

## SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

### Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

### STAFF USE ONLY

Date completed: 05-28-03  
Searcher: Berc. Uy 24954  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

#### Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

#### Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

#### Vendors

\_\_\_\_\_ IG Suite  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:35:43 ; Search time 3485 Seconds  
(without alignments)  
3006.316 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILPSRSGSRGAL.....QNKLGHPAGKRCPLGNRS 360

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q/cg2\_1/USPTO\_spool/US10023586/runat\_22052003\_114850\_28641/app\_query.fasta\_1.519  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10023586.CGN.1.1.2496.0runat\_22052003\_114850\_28641 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pin.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1936	100.0	1082	6	AX458306	Sequence
2	1936	100.0	1334	9	AB065670	Homo sapi
3	1936	100.0	3143	6	AX365511	Sequence
c 4	1936	100.0	164502	2	AC026960	Homo sapi
5	1936	100.0	167084	9	AC092999	Homo sapi
6	1923	99.3	163958	2	AC021773	Homo sapi
7	1806	93.3	1019	9	AF411108	Homo sapi
8	1785	92.2	1020	6	AX458308	Sequence
9	1785	92.2	1076	6	AX375230	Sequence
10	1541	79.6	850	6	AX365514	Sequence
11	1154.5	59.6	209273	2	AC121579	Mus muscu
12	1130	58.4	91107	2	AC116265	Rattus no
c 13	921	47.6	510	6	AX147768	Sequence
c 14	888	45.9	192524	2	AC120742	Rattus no
c 15	732	37.8	681	6	AX244718	Sequence
16	681	35.2	585	6	AX244727	Sequence
17	501	25.9	1074	5	AF069555	Meleagris
18	501	25.9	1977	5	AF031897	Meleagris
19	495	25.6	1163	5	GDP213	G.domesticu
20	460.5	23.8	1086	10	MM0277752	Mus muscu
21	460.5	23.8	215102	2	AL671299	Mus muscu
22	460.5	23.8	215410	10	AC091784	Genomic s
23	453	23.4	123185	2	AP000587	Homo sapi
24	447	23.1	2681	9	BC012104	Homo sapi
25	446.5	23.1	1651	9	HSU40223	Human uridi
26	446.5	23.1	136930	5	AL590151	Zebrafish
27	446	23.0	1177	10	RNP2Y4REC	R.norvegicu
28	445.5	23.0	1232	9	HSGPCRE	H.sapiens g
29	445.5	23.0	1429	6	A62985	Sequence 1
30	445.5	23.0	1429	6	HSP2Y4	Sequence 1
c 31	445.5	23.0	178861	2	AC068231	Homo sapi
32	445.5	23.0	178868	9	AL357752	Human DNA
33	442	22.8	2025	9	HSU07225	Homo sapien
34	442	22.8	2257	9	BC028135	Homo sapi
35	442	22.8	2684	10	RNP2Y4RG	Rattus norv
36	436	22.5	1922	10	RATNGPCP2R	Rat mRNA fo
37	435	22.5	196988	9	AF002761	Homo sapi
38	434.5	22.4	2361	5	ALP2Y8	X.laavis mR
39	420.5	21.7	984	6	AR109625	Sequence
40	420.5	21.7	984	6	E12485	Mouse cDNA
41	420.5	21.7	987	10	AF298899	Mus muscu
42	420.5	21.7	1020	6	AR109898	Sequence
43	420.5	21.7	1976	10	BC027331	Mus muscu
44	420	21.7	1842	6	I34573	Sequence 1
45	420	21.7	1842	6	I36524	Sequence 1

ALIGNMENTS

RESULT 1

Db	601	TATGACATGACGACGACCAAGAGAAATTTTGTATGGCTCTTTTGGCTTACGGCATAGTTCTGCACA	660
Qy	221	LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu	240
Db	661	TTGTCTGGCTTCTTTCCCTCTTGGTCATTTTGGTGTGCTATTACATGATGGTCAGGAG	720
Qy	241	ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis	260
Db	721	CCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT	780
Qy	261	ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis	280
Db	781	CCGACCATCTACTGTGTGTGGCTCTTACCCCTCTGTTTGTGTGCTTCCATATATCAC	840
Qy	281	SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly	300
Db	841	TCGCTCCCTTCACCTCACCATCTGCTTCTTCTCAGGAGTCCAGCTCTTGTATGGC	900
Qy	301	SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro	320
Db	901	AGCCAGTGTGCCCTACAGATATGAGGAGCCCTCTGGTGTGTGAGCAGCTGCCTCAACC	960
Qy	321	SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLysLeuArg	340
Db	961	AGTCTGTACTTCTTCTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAACTGAGG	1020
Qy	341	GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer	360
Db	1021	CAGAACAAAGTTGGTGAGCATCCAGCTGGGAGGAGAGATCCCCAGGGTTGAACAGATCT	1080
RESULT 2			
LOCUS	AB065670	1334 bp	DNA linear PRI 23-JUL-2002
DEFINITION	Homo sapiens gene for seven transmembrane helix receptor,		
ACCESSION	isolate:CBRC7TM_233.		
VERSION	AB065670		
KEYWORDS	AB065670.1 GI:21928614		
SOURCE	Homo sapiens (isolate:CBRC7TM_233) DNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.		
JOURNAL	Genome-wide discovery and analysis of human seven transmembrane		
REFERENCE	unpublished		
AUTHORS	2 (bases 1 to 1334)		
TITLE	Suwa,M.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research		
AUTHORS	Center (CBRC), National Institute of Advanced Industrial Science		
TITLE	and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan		
JOURNAL	(E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,		
REFERENCE	tel:81-3-3599-8080, fax:81-3-3599-8081)		
AUTHORS	This sequence is a seven transmembrane helix receptor candidate		
TITLE	predicted from the whole human genome sequences using our automated		
JOURNAL	system that contains programs of gene		
REFERENCE	finding(GeneDecoder), sequence search,		
AUTHORS	transmembrane helix prediction.		
TITLE	And the sequence is submitted by the collaborative project between		
JOURNAL	[Computational Biology Research Center (CBRC), National Institute		
REFERENCE	of Advanced Industrial Science and Technology (AIST)] and [Genome		
AUTHORS	Science Division, Research Center for Advanced Science and Technology		
TITLE	Technology (RCAST), University of Tokyo].		
JOURNAL	Location/Qualifiers		
REFERENCE	1. .1334		
AUTHORS	/organism="Homo sapiens"		
TITLE	/isolate="CBRC7TM_233"		
JOURNAL	/db_xref="taxon:9606"		
REFERENCE	/chromosome="3"		
AUTHORS	1. .1334		
TITLE	/organism="Homo sapiens"		
JOURNAL	/isolate="CBRC7TM_233"		
REFERENCE	/db_xref="taxon:9606"		
AUTHORS	/chromosome="3"		
TITLE	1. .1334		
JOURNAL	/organism="Homo sapiens"		
REFERENCE	/isolate="CBRC7TM_233"		
AUTHORS	/db_xref="taxon:9606"		
TITLE	/chromosome="3"		

```
misc_feature 201..1134
/product="seven transmembrane helix receptor"
/evidence=not_experimental
BASE COUNT 283 a 368 c 348 g 335 t
ORIGIN
Alignment Scores:
Pred. No.: 4,85e-166 Length: 1334
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-023-586B-2 (1-360) x AB065670 (1-1334)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
|||||
Db 117 ATGCTGTCCATTGTTCTTCCAGGGGAAGCAGAGCGGAGCGCTGTTGGAGCTCTG 176
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40
|||||
Db 177 CTCTGAGGAGCCCTCCCGGACATGGAGAGTGGACATGATACATCACAGGAACA 236
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerile 60
|||||
Db 237 GGTCTCTGCCAGTCTCAGAGAAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 296
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
|||||
Db 297 ATCTTTATCTAGGGTGGCCACTAAATGGCACTGTCTTGTGGCACTCCTGGGCCAAACC 356
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyr 100
|||||
Db 357 AAGCGTGGAGCTGTGCCACCACCTATCTGTGTAACCTGTAGTGGCCGACCTGCTTTAT 416
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpPropheGlyGlu 120
|||||
Db 417 GTGCTATTGGCTTCCTCATCATCACCTACTACTAGATCACAGGTGGCCCTTCGGGGAG 476
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
|||||
Db 477 CTGCTCTGCAAGTGGTGCACTTCCTGTTCTATATCAACCTTTACGGCAGCATCTCGCTG 536
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
|||||
Db 537 CTGACCTGCATCTCTGTCACCACTCTCTAGGTGTGTGCCACCACCTGTCTCGTGGCC 596
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
|||||
Db 597 TACCGGACCGCAGCATGCTGGCTGGGCACCAACCAACCAACCAACCAACCAACCAACCA 656
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
|||||
Db 657 CAGCTGTGCTGCACACTGGCTCTTCCACACGCACTTACATCAATGGCCAGATGATCTGG 716
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
|||||
Db 717 TATGACATGACCAAGCAAGATTTTGATCGCTTTTTCCTACCGCATGTTCTTGACA 776
QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
|||||
Db 777 TTGCTGGCTTTCTTCTCCCTCTGGTCAATTTGGTGTGTATTCACTGATGTCAGGAG 836
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
|||||
Db 837 CCTGATCAAGCCAGAGAGAACCTCATATGAGAGACAGCAACACAGCCAGCCAGGTCCAT 896
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
|||||
Db 897 CCGGACCATCTACTGTGTGTGGCTCTTCAACCTCTGTGTTGTGGCTTCCATATCAC 956
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
|||||
```



\* 36238 36337: gap of 100 bp  
\* 36338 46377: contig of 10040 bp in length  
\* 46378 46477: gap of 100 bp  
\* 46478 59361: contig of 12884 bp in length  
\* 59362 59461: gap of 100 bp  
\* 59462 75221: contig of 15760 bp in length  
\* 75222 75321: gap of 100 bp  
\* 75322 93371: contig of 18050 bp in length  
\* 93372 93471: gap of 100 bp  
\* 93472 117463: contig of 23992 bp in length  
\* 117464 117563: gap of 100 bp  
\* 117564 137115: contig of 20152 bp in length  
\* 137116 137815: gap of 100 bp  
\* 137816 164502: contig of 26687 bp in length.

FEATURES

source

Location/Qualifiers  
1. .164502  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3"  
/clone="RP11-170K4"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. .2917  
/note="assembly\_fragment"  
3018. .6310  
/note="assembly\_fragment"  
6411. .9447  
/note="assembly\_fragment"  
9548. .14470  
/note="assembly\_fragment"  
14571. .20528  
/note="assembly\_fragment"  
20629. .24667  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right"  
24768. .29367  
/note="assembly\_fragment"  
29468. .36237  
/note="assembly\_fragment"  
36338. .46377  
/note="assembly\_fragment"  
46478. .59361  
/note="assembly\_fragment"  
59462. .75221  
/note="assembly\_fragment"  
75322. .93371  
/note="assembly\_fragment"  
93472. .117463  
/note="assembly\_fragment"  
117564. .137115  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:left"  
137816. .164502  
/note="assembly\_fragment"

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

BASE COUNT 48539 a 32273 c 33157 g 49088 t 1405 others  
ORIGIN

Alignment Scores:  
Pred. No.: 2.7e-163 Length: 164502  
Score: 1936.00 Matches: 360  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-023-586b-2 (1-360) x AC026960 (1-164502)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
|||||  
Db 123728 ATGCTGCCATTGCTTCTCCAGGGAAGCAGAGCGGAGCGCTCGTGAGGCTCTG 123669

Qy 21 LeuLeuGluGluAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
|||||  
Db 123668 CTCTGGAGGAGCCCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACA 123609  
  
Qy 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
|||||  
Db 123608 GGTCCTGCCAGTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGATATC 123549  
  
Qy 61 IlePheIleLeuGluLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
|||||  
Db 123548 ATCTTTATCTAGGCGTCCACCTAAATGGACATGCTTGTGGCACCTCTGGGGCAAC 123489  
  
Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
|||||  
Db 123488 AAGCGTGGAGCTGTGCCACCATCTATCTGGTGAACCTGTATGGCCGACCTGCTTTAT 123429  
  
Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120  
|||||  
Db 123428 GTGCTATTGCCCTTCTCATCATCATCTACTACTAGATGACAGGTGGCCCTTCGGGGAG 123369  
  
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
|||||  
Db 123368 CTGCTCTGCAAGCTGGTGCACCTTCTCTGTATATCAACCTTTACGGCAGCATCTGCTG 123309  
  
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
|||||  
Db 123308 CTGACCTGCATCTCTGTGCACCATGTTCTAGGTGTGTGCCACCATGTTGCTGCC 123249  
  
Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValLeu 180  
|||||  
Db 123248 TACCGACCCGAGGATGCTGGCTGGCCAGCACCCACCTGGCCCTGGTGGTCTC 123189  
  
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
|||||  
Db 123188 CAGCTGCTGCCACACTGGCTTCTCCACACGAGTACATCAATGCCAGATGATCTGG 123129  
  
Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
|||||  
Db 123128 TATGACATGACCAAGCAAGAGAAATTTGATCGGCTTTTGGCTACGCGCATAGTTCTGACA 123069  
  
Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
|||||  
Db 123068 TTGCTGGCTTTCTTCCCTTCTTGGTCAATTTTGGTGTGCTATTCACTGATGTCAGGAG 123009  
  
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
|||||  
Db 123008 CCTGATCAAGCCAGAGAGAGAACCTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT 122949  
  
Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
|||||  
Db 122948 CCGGACCATCTACTGTGTGGCTCTTCCACCTCTGTGTGGCTTCCATATATCAC 122889  
  
Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
|||||  
Db 122888 TCGCTCCTTCTACCTCACCATCTGCTTCTTCTTCTCAGGACTGCCAGCTCTTGATGGC 122829  
  
Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnGlnLeuProGlnPro 320  
|||||  
Db 122828 AGCCAGTGTGGCTTACAAGATATGGAGGCTCTGGTGAGTGTGAGCAGCTGCCCTCAACCC 122769  
  
Qy 321 SerProValLeuSerPheLeuGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340  
|||||  
Db 122768 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGAGG 122709  
  
Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
|||||  
Db 122708 CAGAACAAAGTTGGTGAGCATCCAGCTGGGAGGAAGAGATGCCCGAGGTTGAACATCT 122649

RESULT 5  
AC092999  
LOCUS  
DEFINITION

AC092999 167084 bp DNA linear PRI 30-JUL-2002  
Homo sapiens 3 BAC RP11-170K4 (Roswell Park Cancer Institute Human  
BAC Library) complete sequence.



```
/rpt_family="AT_rich"
7595..7717
/rpt_family="MER5A"
complement(7788..8198)
/rpt_family="MLT11"
complement(8536..8591)
/rpt_family="L2"
9634..9936
/rpt_family="Alusx"
complement(10766..10814)
/rpt_family="MIR"
complement(10848..10964)
/rpt_family="MIR"
17399..17431
/rpt_family="(TTAAA)n"
complement(17432..17546)
/rpt_family="MIR"
17637..18073
/rpt_family="Charlielb"
18278..18313
/rpt_family="(TG)n"
21428..21841
/rpt_family="MSPD"
complement(21847..21926)
/rpt_family="LTR33A"
22387..22631
/rpt_family="MLT1G1"
complement(22768..23212)
/rpt_family="Tigger3(Golem)"
complement(23222..23279)
/rpt_family="Tigger3(Golem)"
23311..23437
/rpt_family="MLT1E2"
24105..24225
/rpt_family="MIR"
24226..24517
/rpt_family="MER33"
24518..24629
/rpt_family="MIR"
complement(24639..24724)
/rpt_family="L2"
25420..25994
/rpt_family="L2"
26024..26059
/rpt_family="(TG)n"
26565..27405
/rpt_family="L1MB8"
complement(27789..27855)
/rpt_family="L2"
28628..28778
/rpt_family="MER77"

Alignment Scores:
Pred. No.: 2,75e-163 Length: 167084
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-023-586B-2 (1-360) x AC092999 (1-167084)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyValLeu 20
|||||
Db 160921 ATGCTGTCATTTGCTTCTTCCAGGGGAGCAGAGAGCGGAGCCCTGCTGGAGCTCTG 160980

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40
|||||
Db 160981 CTCCTGGAGGAGCCTCCCGGACATGGAGAAGTGGACATGATACATACAGAGACAA 161040

Qy 41 GlyLeuGlnPheSerGluLysThrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
|||||
Db 161041 GGTCTGCCAGTCTCAGAGAGTACAGCAAGTACCTCTCCCTGGCCTACAGTATC 161100

Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheHisSerTrpGlyGlnThr 80
|||||
Db 161101 ATCTTTATCTAGGGCTGCCACCTAAATGGCACCTCTGTGGCACCTCTCTGGGCCAAACC 161160

Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
|||||
Db 161161 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGTGGCGGACCTGCTTTAT 161220

Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPheGlyGlu 120
|||||
Db 161221 GTGCTATGTCCTTCTCATCATCATCTAGATCAGAGTGGCCCTTCTGGGGAG 161280

Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
|||||
Db 161281 CTGCTGTCAAGCTGGTGCACTTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 161340

Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
|||||
Db 161341 CTGACCTGTCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGGTCC 161400

Qy 161 TyrArgThrArgArgHisAlaTyrLeuGlyThrSerThrThrTyrPheAlaLeuValValLeu 180
|||||
Db 161401 TACCGGACCGCAGGACATGCTGGTGGGCACCACTGGGCGCTGGTGGTCCCTC 161460

Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
|||||
Db 161461 CAGCTGCTGCCACACTGGCTTCTCCACACGAGCTACATCAATGCCAGATGATCTGG 161520

Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
|||||
Db 161521 TATGACATGACAGCCCAAGAGAAATTTGTGCTGGCTTTTGCCTACGCCATAGTTCTGACA 161580

Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
|||||
Db 161581 TTGCTGGCTTCTTCTCCCTCTGGTCAATTTGGTGTGCTATTCATCTGATGGTCAGAG 161640

Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
|||||
Db 161641 CCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGCAACACACGCCCGCAGGTCCAT 161700

Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrHis 280
|||||
Db 161701 CCGGACCATCTACTGTGTGTGGCTCTTCACTCCCTGTGTGTGGCTTCTCCATATCAC 161760

Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
|||||
Db 161761 TCGCTCCTTCTACCTCACCATCTGCTTCTCTCAGAGCTGCCAGCTCTTGTATGGC 161820

Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
|||||
Db 161821 AGCCAGTGTGGCTTACAGATATGAGGCGCTCTGGTCACTGAGAGCTGCCCTCAACCC 161880

Qy 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
|||||
Db 161881 AGTCTGTACTTCTTTCAAGGGGGGCAAAATAGACTCAGGCTCTCCAGAAATCTGAGG 161940

Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
|||||
Db 161941 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 162000

RESULT 6
AC021773 163958 bp DNA linear HTG 10-SEP-2000
LOCUS Homo sapiens clone RP11-12N17, WORKING DRAFT SEQUENCE, 15 unordered
DEFINITION pieces.
AC021773
VERSION AC021773.6 GI:10047750
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```



# REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS

1 (bases 1 to 163958)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome, clone RP11-12N17  
 Unpublished  
 2 (bases 1 to 163958)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,  
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,  
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
 DeArrellano,K., Dekar,K., Domino,M., Doyle,M., Fenestor,J.,  
 Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,  
 Garlyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Landers,T., Lehoczy,J., Levine,P., Lieu,C., Liu,G., Locke,K.,  
 MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., Oliviar,T.M., Peterson,K.,  
 Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
 Roy,A., Santos,R., Severly,P., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
 Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 10, 2000 this sequence version replaced gi:7248967.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center -----  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information -----  
 Center project name: L3302  
 Center clone name: 12\_N17  
 ----- Summary Statistics -----  
 Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 152766 bases at least Q40  
 Consensus quality: 158075 bases at least Q30  
 Consensus quality: 160495 bases at least Q20  
 Insert size: 174000; agarose-fp  
 Insert size: 162558; sum-of-contigs  
 Quality coverage: 4.5 in Q20 bases; agarose-fp  
 Quality coverage: 4.9 in Q20 bases; sum-of-contigs  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 7254: contig of 7254 bp in length  
 \* 7255 7354: gap of 100 bp  
 \* 7355 8612: contig of 1258 bp in length  
 \* 8613 8712: gap of 100 bp  
 \* 8713 11463: contig of 2751 bp in length  
 \* 11464 11563: gap of 100 bp  
 \* 11564 14507: contig of 2944 bp in length  
 \* 14508 14607: gap of 100 bp  
 \* 14608 22452: contig of 7845 bp in length  
 \* 22453 22552: gap of 100 bp  
 \* 22553 35086: contig of 12534 bp in length  
 \* 35087 35186: gap of 100 bp  
 \* 35187 44278: contig of 9092 bp in length  
 \* 44279 44378: gap of 100 bp  
 \* 44379 56786: contig of 12408 bp in length  
 \* 56787 56886: gap of 100 bp

## TITLE JOURNAL COMMENT

\* 56887 65270: contig of 8384 bp in length  
 \* 65271 65370: gap of 100 bp  
 \* 65371 76394: contig of 11024 bp in length  
 \* 76395 76494: gap of 100 bp  
 \* 76495 100320: contig of 23826 bp in length  
 \* 100321 100420: gap of 100 bp  
 \* 100421 113127: contig of 12707 bp in length  
 \* 113128 113227: gap of 100 bp  
 \* 113228 126756: contig of 13529 bp in length  
 \* 126757 126856: gap of 100 bp  
 \* 126857 151061: contig of 24205 bp in length  
 \* 151062 151161: gap of 100 bp  
 \* 151162 163958: contig of 12797 bp in length.

### FEATURES

Location/Qualifiers  
 1..163958  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-12N17"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1..7254  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 7355..8612  
 /note="assembly\_fragment"  
 8713..11463  
 /note="assembly\_fragment"  
 11564..14507  
 /note="assembly\_fragment"  
 14608..22452  
 /note="assembly\_fragment"  
 22553..35086  
 /note="assembly\_fragment"  
 35187..44278  
 /note="assembly\_fragment"  
 44379..56786  
 /note="assembly\_fragment"  
 56887..65270  
 /note="assembly\_fragment"  
 65371..76394  
 /note="assembly\_fragment"  
 76495..100320  
 /note="assembly\_fragment"  
 100421..113127  
 /note="assembly\_fragment"  
 113228..126756  
 /note="assembly\_fragment"  
 126857..151061  
 /note="assembly\_fragment"  
 151162..163958  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right  
 BASE COUNT 49230 a 32369 c 32422 g 48536 t 1401 others  
 ORIGIN

### Alignment Scores:

Pred. NO.: 4.05e-162 Length: 163958  
 Score: 1923.00 Matches: 360  
 Percent Similarity: 99.72% Conservative: 0  
 Best Local Similarity: 99.72% Mismatches: 0  
 Query Match: 99.33% Indels: 1  
 DB: 2 Gaps: 0

US-10-023-586B-2 (1-360) x AC021773 (1-163958)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
 |||||  
 Db 98103 ATGCTGTCCATTTTCTCTCCAGGGGAGCAGAGCGGAGCGTGGAGCTCTG 98162  
 |||||  
 Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
 |||||  
 Db 98163 CTCCTGGAGGAGCGCTCCCGGACATGGAGAGGTGGACATCAATACATCAGAGNACAA 98222



```
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
|||
Db 98223 GGTCCTGCCAGTCTCCAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 98282

QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
|||
Db 98283 ATCTTTATCTAGGCTGGCCTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 98342

QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
|||
Db 98343 AAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 98402

QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpPropheGlyGlu 120
|||
Db 98403 GTGCTATTGCCCTTCCTCATCATCACTACTAGTACAGAGTGGCCCTTCGGGGAG 98462

QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
|||
Db 98463 CTGCTCTGAAGCTGGTGCACTTCCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTG 98522

QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
|||
Db 98523 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTCTCGCTGCC 98582

QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
|||
Db 98583 TACCGGACCGCAGGATGCCTGGCTGGCCAGCAGCACCACTGGCCCTGGTGGTCTTC 98642

QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
|||
Db 98643 CAGCTGTGCCACACAGTGGCTTCTCCACAGCACTATCAATGCCAGATGATCTGG 98702

QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
|||
Db 98703 TATGACATGACCAAGCAAGAAATTTGTATGGCTTTTTCCTACCGCATAGTCTTGACA 98762

QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
|||
Db 98763 TTGCTGTGGCTTTCTTCCCTCCTTGTGTATTTTGGTGTGTATTCATGATGTCAGGAG 98822

QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
|||
Db 98823 CCTGATCAAGCCAGAGGAGAACTCATGAGGAGCAACACAGCCGACGAGTCCAT 98882

QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
|||
Db 98883 CCGGACCATCTACTGTGTGGCTCTTTCACCCCTCTGTGTGGCTTCCATATCAC 98942

QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
|||
Db 98943 TCGCTCTCTTACCTCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCC 99002

QY 301 Ser-GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPr 320
|||
Db 99003 AGCCCAAGTGTGGCTCAAGATATGAGGCTCTGTGTAGTGTGAGCAGTGCCTCAACC 99062

QY 320 oSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLysLeuAr 340
|||
Db 99063 CAGTCTGTACTTCTTTCAGGGGGGCAAAATAGATCAGGCTCTCCAGAACTGAG 99122

QY 340 gGlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgse 360
|||
Db 99123 GCAGAACAAAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGACAGATC 99182

QY 360 r 360
|||
Db 99183 T 99183

RESULT 7
AF411108
LOCUS
DEFINITION Homo sapiens GPR79 pseudogene, complete sequence.
1019 bp DNA linear PRI 01-NOV-2001
```

```
AF411108
AF411108.1 GI:16566321
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1019)
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhicko,O.,
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
TITLE Discovery and mapping of ten novel G protein-coupled receptor genes
JOURNAL Gene 275 (1), 83-91 (2001)
MEDLINE 21458557
PUBMED 11574155
REFERENCE 2 (bases 1 to 1019)
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhicko,O.,
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2001) Department of Pharmacology, University of
Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada
FEATURES
source
1..1019
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1019
/gene="GPR79"
/pseudo
BASE COUNT 200 a 308 c 252 g 259 t
ORIGIN
Alignment Scores:
Pred. No.: 2,06e-154 Length: 1019
Score: 1806.00 Matches: 339
Percent Similarity: 99.71% Conservative: 0
Best Local Similarity: 99.71% Mismatches: 0
Query Match: 93.29% Indels: 93 1
DB: 9 Gaps: 0
US-10-023-586B-2 (1-360) x AF411108 (1-1019)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
|||
Db 1 ATGCTCTCCATTTTGTCTTCCAGGGAAGCAGAGCGGAGCGTCTGGAGCTCTG 60
|||
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
|||
Db 61 CTCTCTGGAGGAGCCCTCCCGGACATGGAGAGTGGACATGAATACATCACAGGAACA 120
|||
QY 41 GlyLeuCysGlnPheSerGluTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
|||
Db 121 GGTCCTGCCAGTCTCCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
|||
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
|||
Db 181 ATCTTTATCTAGGCTGGCCTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240
|||
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
|||
Db 241 AAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300
|||
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpPropheGlyGlu 120
|||
Db 301 GTGCTATTGCCCTTCCTCATCATCACTACTAGTACAGAGTGGCCCTTCGGGGAG 360
|||
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
|||
Db 361 CTGCTCTGAAGCTGGTGCACTTCCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTG 420
|||
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
|||
Db 421 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTCTCGCTGCC 480
|||
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
```



Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R., Elliot, V.S., Ramkumar, J., Baughn, M.R., Kallik, D.A., Wadia, N.K., Hafalia, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Pollicky, J.L., Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.  
G-protein coupled receptors  
Patent: WO 0210387-A 20 07-FEB-2002;  
Incyte Genomics, Inc. (US)

FEATURES  
source  
1..1076  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7474806CB1"

BASE COUNT 211 a 318 c 274 g 273 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.77e-152 Length: 1076  
Score: 1785.00 Matches: 337  
Percent Similarity: 99.12% Conservative: 1  
Best Local Similarity: 98.83% Mismatches: 1  
Query Match: 92.20% Indels: 2  
DB: 6 Gaps: 0

US-10-023-586B-2 (1-360) x AX375230 (1-1076)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgGlySerArgGlyAlaLeu 20  
Db 57 ATGCTGTCATTTTCCTTCCTCAGGGAAGACAGAGCGGAGCCCTGCGGAGCTG 116  
Qy 21 LeuLeuGluGlyAlaSerArgGlyMetGluGlyValAspMetAsnThrSerGlnGln 40  
Db 117 CTCCTGGAGGAGCTCCCGGACATGGAGAGGTGACATGATACATACAGACAA 176  
Qy 41 GlyLeuGlnPheSerGluGlyTyrGlnValTyrLeuSerLeuAlaTyrSerile 60  
Db 177 GGTCTCTGCCAGTTCCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATC 236  
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrpGlyGlnThr 80  
Db 237 ATCTTTATCTAGGCGGCCCTAAATGGCACTGCTGTGGCACTCTGGGGCAACCC 296  
Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
Db 297 AAGCGCTGGAGCTGTCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTAT 356  
Qy 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpPheGlyGlu 120  
Db 357 GTGCTATTGGCTTCTCCTCATCATCCTACTACTAGTACAGGTGCGGCTTCCGGGAG 416  
Qy 121 LeuLeuGlyLeuValHisPheLeuPheThrIleAsnLeuThrGlySerIleLeuLeu 140  
Db 417 CTGCTCTGCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGGAGCATCTGCTG 476  
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
Db 477 CTGACCTGCATCTCTGTCACCACTCTCTAGTGTGTGGCACCCACATGTGTCTGCC 536  
Qy 161 TyrArgThrArgGlyHisAlaThrPheLeuGlyThrSerThrThrPheAlaLeuValLeu 180  
Db 537 TACCGGACCCGACGATGCTGCTGGCACCAAGTGTGGTGGTGGTGGTGGTGGTGGT 596  
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspThrIleAsnGlyGlnMetIleTrp 200  
Db 597 CAGTGTGTCGCCACACGAGGCTTCTCCACACGAGTACATCAATGCCAGATGATCTGG 656  
Qy 201 TyrAspMetThrSerGlnGlnAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
Db 657 TATCATGACGACGACGAGAAATTTGATCGGCTTTTGTGCTACGCAATGCTCTGACA 716  
Qy 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1..850  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 165 a 262 c 201 g 222 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.68e-130 Length: 850  
Score: 1541.00 Matches: 283  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 79.60% Indels: 0  
DB: 6 Gaps: 0

US-10-023-586B-2 (1-360) x AX365514 (1-850)

Qy 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSerGluLys 48  
Db 1 ATGAGAGAGGTGGACATGATATACATACAGGACAAAGGTCTCTGCCAGTCTCAGAGAG 60  
Qy 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 68  
Db 61 TACAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTTATCTTAGGCTGCCACTA 120  
Qy 69 AsnGlyThrValLeuThrPheSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
Db 121 AATGGCACTGTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180  
Qy 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuIle 108  
Db 181 TATCTGTGAGACCTGATGGTGGGCGACCTGCTTTATGTGCTTATTTGCTTCTCTCATC 240  
Qy 109 ThrTyrSerLeuAspArgTrpPheGlyGluLeuLeuCysLysLeuValHisPhe 128

1717 TTGCTGGCTTTCTTTCCCTCCTTGGTGCATTTTGGTGTGCTATTCTACGTGATGGTCCAGG 776  
240 luProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260  
777 AGCTGATCAGCAGGAGAGACCTCATGAGGACAGCAACAGCCAGCCAGGTC 836  
260 isProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280  
837 ATCCGGACCATCTACTGCTGGTGGCTCTTCCACCTCTGTTTGTGGCTTCCATATC 896  
280 isSerLeuLeuProHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300  
897 ACTGCTCTCTTACCTCAGCATCTGCTTCTGCTTCTCAGGACTGCCAGCTTGTATG 956  
300 lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluLeuProGlnP 320  
957 GCACCCAGTGTGGCTACAAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAAC 1016  
320 roSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeu 339  
1017 CCAGTCTCTTCTTCAAGGGGGGCAAAATAGATCAGTCCAGTCCCTCCAGAACTG 1075  
RESULT 10  
AX365514  
LOCUS  
DEFINITION Sequence 4 from Patent WO0185764.  
ACCESSION AX365514  
VERSION AX365514.1 GI:18697083  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1  
AUTHORS Ramakrishnan, S.  
TITLE Regulation of human p2y-like g protein-coupled receptor  
JOURNAL Regulation of human p2y-like g protein-coupled receptor  
Bayer Aktiengesellschaft (DE)  
FEATURES  
source  
1..850  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

```
Db 241 ACCTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 300
Qy 129 LeuPheThrIleAsnLeuTyrglySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCTATCTCTGTGCACAG 360
Qy 149 PheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrglyThrArgArgHisalaTrp 168
Db 361 TTCTCTAGGTGTGTGCCACCCACTGTGTTCGTGCTGCTACCGACCCGACGAGCATGCTGTG 420
Qy 169 LeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuLeuProThrLeuAlaPhe 188
Db 421 CTGGGCACGACGACCATCTGGGCCCTGGTGGTCTCCAGCTGCTGCCACACTGGCCCTTC 480
Qy 189 SerHisThrAspTyrlleAsnGlyGlnMetIleTrpTyrglyMetThrSerGlnGluAsn 208
Db 481 TCCACACGGACTACATCAATGCGCCAGATGATCTGATGACATGACATGACGACGACGAGAA 540
Qy 209 PheAspArgLeuPheAlaTyrglyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228
Db 541 TTGTATCGGCTTTTGGCTACGGCAGTATGCTGACATGCTGCTGCTTTCTTCCCTCCTT 600
Qy 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248
Db 601 GGTCTATTTGGTGTCTTACTGATGCTGTCAGGAGCTGATCAGACGACGAGAGAACT 660
Qy 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268
Db 661 CATGAGGACGACCAACAGCCGAGCCAGGTCCTATCGGACCATCTACTGTTGTGTGG 720
Qy 269 ProLeuHisProLeuPheCysAlaLeuProTyrglyHisSerLeuLeuLeuProHisHisLeu 288
Db 721 CCTCTTCAACCTCTCTTTGTCCTTCCATATCACTGCTGCTTCTTCTACCTACCATCTG 780
Qy 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 781 CTTTCTGCTTCTCAGGACTGCCACCTTGTATGACGACCATCTGCTGCTGCTGCTGCTG 840
Qy 309 GluAlaSer 311
Db 841 GAGGCTCT 849

RESULT 11
AC121579
LOCUS Mus musculus chromosome UNK clone RP23-255H8, WORKING DRAFT ;
DEFINITION AC121579 209273 bp DNA linear HTG 20-MAY-2002
SEQUENCE, 9 unordered pieces.
ACCESSION AC121579
VERSION AC121579.1 GI:20986651
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 209273)
JOURNAL McPherson,J.D. and Waterston,R.H.
REFERENCE The sequence of Mus musculus clone
2 (bases 1 to 209273)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M.BA025H08
----- Summary Statistics -----
```

```
Sequencing vector: M13; 0%
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 205908 bases at least Q40
Consensus quality: 206285 bases at least Q30
Consensus quality: 206424 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 208741; sum-of-contigs
Quality coverage: 15.45 in Q20 bases; agarose-fp
Quality coverage: 12.20 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1957: contig of 1957 bp in length
* 1958 2057: gap of unknown length
* 2058 7480: contig of 5423 bp in length
* 7481 7580: gap of unknown length
* 7581 21224: contig of 13644 bp in length
* 21225 21324: gap of unknown length
* 21325 40793: contig of 19469 bp in length
* 40794 40893: gap of unknown length
* 40894 73232: contig of 32339 bp in length
* 73233 73332: gap of unknown length
* 73333 113562: contig of 40230 bp in length
* 113563 113662: gap of unknown length
* 113663 165371: contig of 51709 bp in length
* 165372 165471: gap of unknown length
* 165472 208277: contig of 42806 bp in length
* 208278 208377: gap of unknown length
* 208378 209273: contig of 896 bp in length.
*
* Location/Qualifiers
* 1. 209273
* /organism="Mus musculus"
* /db_xref="taxon:10090"
* /chromosome="UNK"
* /clone="RP23-255H8"
* 1. 1957
* /note="assembly_name:Contig18"
* 2058. 7480
* /note="assembly_name:Contig19"
* 7581. 21224
* /note="assembly_name:Contig20"
* 21325. 40793
* /note="assembly_name:Contig21"
* 40894. 73232
* /note="assembly_name:Contig22"
* 73333. 113562
* /note="assembly_name:Contig23"
* 113663. 165371
* /note="assembly_name:Contig24"
* 165472. 208277
* /note="assembly_name:Contig25"
* 208378. 209273
* /note="assembly_name:Contig16"
*
BASE COUNT 59147 a 42827 c 43921 g 62577 t 801 others
ORIGIN

Alignment Scores:
Pred. No.: 2,46e-92 Length: 209273
Score: 1154.50 Matches: 236
Percent Similarity: 74.37% Conservative: 31
Best Local Similarity: 65.74% Mismatches: 83
Query Match: 59.63% Indels: 10
DB: 2 Gaps: 2
```

US-10-023-586b-2 (1-360) x AC121579 (1-209273)

```
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerArgGlyAlaLeu 20
Db 37890 ATGCTGGCTTTATTTCTTCAGAGAACG-----AGAGTCGACGC 37934
QY 21 LeuLeuGluGly-AlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40
Db 37935 CATCTTCGAGGAGCCTCTTTAGGGATGGAGAGCTAGATATCATCTCTCAAGAGGTGC 37994
QY 40 nGlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 37995 TGGCGCTCGCCACTTCTCAGAGAGCTACAAGCAAGCTACCTATCATACCTACAGTCT 38054
QY 60 eilePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrpGlyGln 80
Db 38055 CATCTTTATGCTAGGATGCGCTCAATGGGGCTGTCTATGGCTCTCTGGTGCACAAAC 38114
QY 80 rLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTy 100
Db 38115 CAAGCGCTGAGCTGTGCCACCATCTACCTGATGAACCTGGTGGCTGACCTGCTTTA 38174
QY 100 rValLeu---LeuProPheLeuIleThrTyrSerLeuAspArgTrpPheGln 119
Db 38175 TATGCTAAGCTGCGCTTGTCTATCATCATCTACCTACGGGGTGGGTGACACCTGGCTTTGG 38234
QY 119 YGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLe 139
Db 38235 GGAGATGCTGTGCAGGTGTGGGTCTCTTCTTCTACACCACTCTACAGCAGCATCCT 38294
QY 139 uLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLe 159
Db 38295 GCTGCTGACCTGCTCTCTACACCGCTCTCTAGGAGTGTGCACCTCTGAGCTCACT 38354
QY 159 uProTyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuVal 179
Db 38355 GTCTACCGGAGCTACCGGCGCCGCTGGCAGGAGCTGCGCCACCTGGACCTGGTGTAGT 38414
QY 179 lLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIle 199
Db 38415 TCTCCAATTGCTGCCGACTCTGCTCTCTCCACACAGACTATGTAATGGCCAGGTGAT 38474
QY 199 eTrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLe 219
Db 38475 CTGCTATGACACAGCCGCGGAGGAGTGTGTATCTTCTGCGCTACAGCATAGTACT 38534
QY 219 uThrLeuSerGlyPhe--LeuSerLeuGlyHisPheGlyValLeuPheThrAspGly 238
Db 38535 GAGCTTATCTGGGTTTCTCTCCCTCTCTGTATCATCTTCTGGGTGCTACTCTCTGATGGT 38594
QY 239 GlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGln 258
Db 38595 TAGGAGCTGACAAATCCAGTGGAGGACCTCCCAACGACAGCCACACGCCGTGCCAA 38654
QY 259 ValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuPro 278
Db 38655 GTCTATTGCGACCATCTGCTGTGTGTGGCTCTCTCACACTCTGTTTGTGCCCTTCCA 38714
QY 279 TyrHisSerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeu 298
Db 38715 CATTCGCCGATCTCTTCTATGTATCTCCGCTCTCTGAAGTCACAGGATGCCAGCTCTT 38774
QY 299 AspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnLeuPro 318
Db 38775 GTGGCAGCCAGCATGCGCTACAGGCTCTAGGCTCTGTGAAGCATGAGCAGCTGCTCT 38834
QY 319 GlnProSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLys 338
Db 38835 CAATCCAGTCTGCTACTTC--TTGTACAGGGCAACACAGAGCCCACTCTCCAGAG 38892
QY 339 LeuArgGlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysPro 355
Db 38893 CTGAGACAAACAAAGGTGGGTGAGCACCCCTCTGTAAGGAGGACCT 38943
```

## RESULT 12

AC116265/c

LOCUS

DEFINITION

AC116265

AC116265

VERSION

HTG: HTGS\_PHASE1.

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 91107)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T.,

Barbata, J., Benton, J., Bimble, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Chen, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,

Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,

Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louseghed, H.,

Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,

Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,

Moser, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,

Muganyizi, N., Nickerson, E., Nwokoko, S., Ogih, M., Okwuonu, G.,

Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,

Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G.,

Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,

Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,

Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,

Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,

Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,

Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,

Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,

Wu, C., Wu, Y., Wu, X.F., Zhou, J., Zorrilla, S., Nelson, D.,

Weinstock, G., and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 91107)

Worley, K.C.

Direct Submission

Submitted (26-MAR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 91107)

Worley, K.C.

Direct Submission

Submitted (18-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 14, 2002 this sequence version replaced gi:19718459.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GTRS  
 Center clone name: CH230-18307  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 36430 bases at least Q40  
 Consensus quality: 38793 bases at least Q30  
 Consensus quality: 40859 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 46 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1
* 1061: contig of 1061 bp in length
* 1161: gap of unknown length
* 1162: contig of 1556 bp in length
* 2717: gap of unknown length
* 2818: gap of unknown length
* 4467: contig of 1650 bp in length
* 4468: gap of unknown length
* 4568: contig of 1529 bp in length
* 6097: gap of unknown length
* 6197: gap of unknown length
* 7437: gap of unknown length
* 7537: gap of unknown length
* 8768: gap of unknown length
* 8868: gap of unknown length
* 10134: contig of 1266 bp in length
* 10233: gap of unknown length
* 10234: contig of 1287 bp in length
* 11521: gap of unknown length
* 11621: gap of unknown length
* 11622: gap of unknown length
* 13037: gap of unknown length
* 13137: contig of 1091 bp in length
* 14228: gap of unknown length
* 14328: gap of unknown length
* 15362: gap of unknown length
* 15461: gap of unknown length
* 16999: gap of unknown length
* 17099: gap of unknown length
* 18220: gap of unknown length
* 18320: gap of unknown length
* 19564: gap of unknown length
* 19664: gap of unknown length
* 21323: contig of 1660 bp in length
* 21424: gap of unknown length
* 22610: contig of 1186 bp in length
* 22710: gap of unknown length
* 24221: gap of unknown length
* 24321: gap of unknown length
* 25466: contig of 1146 bp in length
* 25566: gap of unknown length
* 25567: contig of 1731 bp in length
* 27298: gap of unknown length
* 27398: gap of unknown length
* 28794: gap of unknown length
* 28894: gap of unknown length
* 30289: contig of 1396 bp in length
* 30290: gap of unknown length
* 30390: contig of 1357 bp in length
* 31747: gap of unknown length
* 31846: gap of unknown length
* 33603: contig of 1757 bp in length
* 33703: gap of unknown length
* 33704: contig of 1558 bp in length
* 35261: gap of unknown length
* 35361: gap of unknown length
* 37315: contig of 1954 bp in length
* 37316: gap of unknown length

```

```

* 37416 38787: contig of 1372 bp in length
* 38788 38887: gap of unknown length
* 38888 40559: contig of 1672 bp in length
* 40560 40659: gap of unknown length
* 40660 42655: contig of 1996 bp in length
* 42656 42755: gap of unknown length
* 42756 44162: contig of 1407 bp in length
* 44163 44262: gap of unknown length
* 44263 45600: contig of 2238 bp in length
* 45601 46500: gap of unknown length
* 46501 48572: contig of 1972 bp in length
* 48573 48672: gap of unknown length
* 48673 50544: contig of 1872 bp in length
* 50545 50644: gap of unknown length
* 50645 51883: contig of 1239 bp in length
* 51884 51983: gap of unknown length
* 51984 54568: contig of 2585 bp in length
* 54569 54668: gap of unknown length
* 54669 56245: contig of 1577 bp in length
* 56246 56345: gap of unknown length
* 56346 58742: contig of 2397 bp in length
* 58743 58842: gap of unknown length
* 58843 61978: contig of 3136 bp in length
* 61979 62078: gap of unknown length
* 62079 63933: contig of 1855 bp in length
* 63934 64033: gap of unknown length
* 64034 65836: contig of 1803 bp in length
* 65837 65936: gap of unknown length
* 65937 68259: contig of 2323 bp in length
* 68260 68359: gap of unknown length
* 68360 71042: contig of 2583 bp in length
* 70943 71043: contig of 3535 bp in length
* 71044 74677: gap of unknown length
* 74678 76364: contig of 1687 bp in length
* 76365 76464: gap of unknown length
* 76465 80009: contig of 3545 bp in length
* 80010 80109: gap of unknown length
* 80110 85962: contig of 5853 bp in length
* 85963 86062: gap of unknown length
* 86063 91107: contig of 5045 bp in length.;
FEATURES             Location/Qualifiers
     source            1..91107
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-18307"
BASE COUNT    25705 a 18632 c 18796 g 23400 t 4574 others
ORIGIN
Alignment Scores:
Pred. No.:      1.37e-90      Length:      91107
Score:          1130.00      Matches:    233
Percent Similarity: 74.72%      Conservative: 36
Best Local Similarity: 64.72%      Mismatches: 79
Query Match:     58.37%      Indels:     13
DB:              2              Gaps:        4
US-10-023-586B-2 (1-360) x AC116265 (1-91107)
Qy      1 MetLeuSerIleLeuLeuProSerArgGlySerArgGlySerArgGlyAlaLeu 20
      ||||| :||||| ||| ||| |||||
Db 78493 ATGCTGTGCGTTTATTTCCTTACAGGAGCAGCAAGT-----AGAGGCCAT--- 78446
Qy      21 LeuLeu-GluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluG1 40
      ||| ||| ||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 78445 ---CTTTGAGGAAGCCTCCGTTGAGATGGAGAGCTACACGTCTCAGAGATGA 78389
Qy      40 nGlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerI1 60
      :||| ||| ||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 78388 AGGCTTCTGCCACTTCTCGGAGAGCTACAAAGCAAGCTTACCTATCGCTAACCTACAGTCT 78329
Qy      60 eilePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnTh 80
      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```



Db 78328 CATCTTCATGCTAGGAATGCCCTCAATAGCGCCGCTTATGGCTCTCTCTGGTGCACAAAC 78269  
 QY 80 rLysArgTprSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTy 100  
 Db 78268 CAAGCCCTGGAGCTGTACCACTATCTACCTGTATGACCTGTATGGTCCGACCTGCTTTA 78209  
 QY 100 rValLeu---LeuProPheLeuLeuLeuThrTyrSerLeuAspArgTprProPheG1 119  
 Db 78208 CATGCTGAGTCTGCCCTCTCTCATCATCACTACGCGCGGGTGACACCTGGCTCTTTGG 78149  
 QY 119 yGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLe 139  
 Db 78148 GGAGAGCGCTGTGAGCGTGTGGCTTCTTCTACACCACTCTCTACAGCAGCATCTCT 78089  
 QY 139 uLeuLeuThrCysLysSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLe 159  
 Db 78088 GCTGCTGAGCTGATCTCTCTATACCGCTTCTAGAGGTGCGCACCCCTGAGCTCACT 78029  
 QY 159 uProTyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValVa 179  
 Db 78028 GTCCCTCCGACCCACCGCATACCTGGCAAGCAGCGCCACCATCTGGACCTGGTAGT 77969  
 QY 179 lLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetII 199  
 Db 77968 CGTCCAAATCTGCTACCGACTCTGTCTCTCCACACAGACTATGTAATGGCCACCTGAT 77909  
 QY 199 eTrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLe 219  
 Db 77908 CTGCTACGACACAGCCAGCCAGAGATTTTCGATCATTTCTCGCGTGTCTCTCTGATGT 77849  
 QY 219 uThrLeuSerGlyPhe--LeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGly 238  
 Db 77848 GACACTATCTGGGTGTCT 77789  
 QY 239 GlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGln 258  
 Db 77788 CAGGAGCTGACAAATCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 77729  
 QY 259 ValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuPro 278  
 Db 77728 GTCTATCCGACCATCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 77669  
 QY 279 TyrHisSerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeu 298  
 Db 77668 CATGGCCGGTCT 77609  
 QY 299 AspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGlyCysGluGlnLeuPro 318  
 Db 77608 GTTGGCAGCAGCATGGCTTACAGATCTGTAGGCTCTGTAGGCTCTGTAGGCTCTGT 77549  
 QY 319 GlnProSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLys 338  
 Db 77548 CAACCCGGTCTGTACTTC--TTGTACAGGGCAACACAGAGTCCAACTCTCTCCAGAG 77491  
 QY 339 LeuArgGlnAsnLysLeuGlyGlyGlyHisProAla-----GlyArgLysArg 353  
 Db 77490 CTGAGACAAACAGGTGGGTGAGCACCCTCACTAGTGTGGAAGGAGCAGCAGTAGA 77437  
 RESULT 13  
 AX147768  
 LOCUS AX147768 510 bp DNA linear PAT 08-JUN-2001  
 DEFINITION Sequence 13 from Patent WO0136473.  
 ACCESSION AX147768  
 VERSION AX147768.1 GI:14346805  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 510)  
 Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,  
 Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,  
 Sejlitz,T. and Huff,R.M.

TITLE Novel g protein-coupled receptors  
 JOURNAL Patent: WO 0136473-A 13 25-MAY-2001;  
 PHARMACIA & UPJOHN COMPANY (US)

FEATURES  
 source Location/Qualifiers  
 1..510

BASE COUNT 92 a 156 c 126 g 136 t  
 ORIGIN /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

Alignment Scores:  
 Pred. No.: 1.32e-74 Length: 510  
 Score: 921.00 Matches: 170  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 47.57% Indels: 0  
 DB: 6 Gaps: 0

US-10-023-586B-2 (1-360) x AX147768 (1-510)

QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102  
 Db 1 TGGAGCTGTGCCACACCTATCTGTGAGACCTGATGTGGCCGACCTGCTTTATGCTGA 60  
 QY 103 LeuProPheLeuLeuIleThrTyrSerLeuAspArgTprProPheGlyGluLeuLeu 122  
 Db 61 TTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120  
 QY 123 CysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThr 142  
 Db 121 TGAAGCTGTGTGACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTGACC 180  
 QY 143 CysLysSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg 162  
 Db 181 TGCATCTCTGTGACCACTCTCTAGGTGTGTGCCACCACTGTGTTGCTGCGCTACCGG 240  
 QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeu 182  
 Db 241 ACCCGCAGGATGCTGGCTGGCCACCACTACCTGGCCCTGGTGGTCTCCAGCTG 300  
 QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAsp 202  
 Db 301 CTGCCACACTGGCTTCTCTCCACACGACTACATCAATGGCCAGATGATCTGGTATGAC 360  
 QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer 222  
 Db 361 ATGACCAAGCAAGAGATTTTGTAGCGGCTTTTCCCTACGCGATGATCTGACATGTCT 420  
 QY 223 GlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAsp 242  
 Db 421 GGCTTTCTTCCCTCTTGGTCAATTTGGTGTGCTATTCTACTGATGTCAGGAGCTGAT 480  
 QY 243 GlnAlaArgGlyGluProHisGluAspArg 252  
 Db 481 CAAGCAGGAGGAGAACCTCATGAGGACAGG 510

RESULT 14  
 AC120742/c  
 LOCUS AC120742 192524 bp DNA linear HTG 24-JUL-2002  
 DEFINITION Rattus norvegicus clone CH230-15014, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 56 unordered pieces.  
 ACCESSION AC120742  
 VERSION AC120742.3 GI:21908162  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Rattus norvegicus.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 192524)  
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barberia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,

Bouck, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.P., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Delucchi, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
 Honsi, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozad, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, K., Martindale, A., Martinez, E.,  
 Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newtson, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,  
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Williams, G., Williams, A., Wleczyk, R., Woodson, S., Worley, K.,  
 Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G. and Gibbs, R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 192524)  
 Worley, K.C.

Direct Submission  
 Submitted (09-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 192524)  
 Worley, K.C.

Direct Submission  
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 19, 2002 this sequence version replaced gi:20564367.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GXVJ  
 Center clone name: CH230-15014  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 137507 bases at least Q40  
 Consensus quality: 143272 bases at least Q30  
 Consensus quality: 147627 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 56 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1027: contig of 1027 bp in length  
 1127: gap of unknown length  
 1128: contig of 1024 bp in length  
 2151: gap of unknown length  
 2251: gap of unknown length  
 3522: contig of 1271 bp in length  
 3523: gap of unknown length  
 3623: contig of 1443 bp in length  
 5165: gap of unknown length  
 5166: contig of 1695 bp in length  
 6860: gap of unknown length  
 7995: contig of 1035 bp in length  
 8095: gap of unknown length  
 9381: contig of 1286 bp in length  
 9481: gap of unknown length  
 10544: contig of 1063 bp in length  
 10644: gap of unknown length  
 12080: contig of 1436 bp in length  
 12180: gap of unknown length  
 13574: contig of 1394 bp in length  
 13674: gap of unknown length  
 14766: contig of 1092 bp in length  
 14866: gap of unknown length  
 16102: contig of 1236 bp in length  
 16202: gap of unknown length  
 17487: contig of 1285 bp in length  
 17587: gap of unknown length  
 18875: contig of 1288 bp in length  
 18975: gap of unknown length  
 20257: contig of 1282 bp in length  
 20357: gap of unknown length  
 21924: contig of 1567 bp in length  
 22024: gap of unknown length  
 23072: contig of 1048 bp in length  
 23172: gap of unknown length  
 24993: contig of 1827 bp in length  
 25099: gap of unknown length  
 26426: contig of 1327 bp in length  
 26526: gap of unknown length  
 27831: contig of 1305 bp in length  
 27931: gap of unknown length  
 29815: contig of 1885 bp in length  
 29916: gap of unknown length  
 31690: contig of 1774 bp in length  
 31791: gap of unknown length  
 34578: contig of 2788 bp in length  
 34678: gap of unknown length  
 36533: contig of 1855 bp in length  
 36633: gap of unknown length  
 38324: contig of 1691 bp in length  
 38424: gap of unknown length  
 40498: contig of 2074 bp in length  
 40598: gap of unknown length  
 43132: contig of 2534 bp in length  
 43232: gap of unknown length  
 46098: contig of 2866 bp in length  
 46198: gap of unknown length  
 48303: contig of 2105 bp in length  
 48403: gap of unknown length  
 51105: contig of 2703 bp in length  
 51206: gap of unknown length  
 54222: contig of 3016 bp in length  
 54322: gap of unknown length  
 56846: contig of 2524 bp in length  
 56946: gap of unknown length  
 60999: contig of 4053 bp in length  
 61099: gap of unknown length  
 65048: contig of 3949 bp in length  
 65148: gap of unknown length



\* 65149 67208: contig of 2060 bp in length  
\* 67209 67308: gap of unknown length  
\* 67309 70014: contig of 2706 bp in length  
\* 70015 70114: gap of unknown length  
\* 70115 75273: contig of 5159 bp in length  
\* 75274 75373: gap of unknown length  
\* 75374 78558: contig of 3185 bp in length  
\* 78559 78659: gap of unknown length  
\* 78659 81049: contig of 2391 bp in length  
\* 81050 81149: gap of unknown length  
\* 81150 84928: contig of 3779 bp in length  
\* 84929 85028: gap of unknown length  
\* 85029 89226: contig of 4198 bp in length  
\* 89227 89326: gap of unknown length  
\* 89327 94002: contig of 4676 bp in length  
\* 94003 94102: gap of unknown length  
\* 94103 98447: contig of 4345 bp in length  
\* 98448 98547: gap of unknown length  
\* 98548 102392: contig of 3745 bp in length  
\* 102393 102392: gap of unknown length  
\* 102393 106997: contig of 4605 bp in length  
\* 106998 107097: gap of unknown length  
\* 107098 111457: contig of 4360 bp in length  
\* 111458 111557: gap of unknown length  
\* 111558 115579: contig of 4022 bp in length  
\* 115580 115679: gap of unknown length  
\* 115680 119962: contig of 4283 bp in length  
\* 119963 120062: gap of unknown length  
\* 120063 126366: contig of 6304 bp in length  
\* 126367 126466: gap of unknown length  
\* 126467 133796: contig of 7330 bp in length  
\* 133797 133896: gap of unknown length  
\* 133897 142152: contig of 8256 bp in length  
\* 142153 142252: gap of unknown length  
\* 142253 151202: contig of 8950 bp in length  
\* 151203 151302: gap of unknown length  
\* 151303 158779: contig of 7477 bp in length  
\* 158780 158879: gap of unknown length  
\* 158880 168094: contig of 9215 bp in length

## Alignment Scores:

Pred. No.: 3,128-68 Length: 192524  
Score: 888.00 Matches: 176  
Percent Similarity: 75.56% Conservative: 28  
Best Local Similarity: 65.19% Mismatches: 57  
Query Match: 45.87% Indels: 9  
DB: 2 Gaps: 3

US-10-023-586B-2 (1-360) x AC120742 (1-192524)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20  
DB 159732 ATGCTGGGTTTATTTCTTACAGGAGCAGCAGAGT-----AGAGCCAT--- 159685  
QY 21 LeuLeu-GluGlyAlaSerArgMetGluLysValAspMetAsnThrSerGlnGlu 40  
DB 159684 ---CTTTGAGGAGCCCTCGCTTGCAGTGGAGAGCTAGACGTAATGCTCTCAGAGGATGA 159628  
QY 40 nGlyLeuCysGlnPheSerGluLysTrpLysGlnValTyLeuSerLeuAlaTySer 60  
DB 159627 AGGCTTCTGCCACTTCTCGGAGAGCTACAGCAAGTCTACCTACCTACCTACAGTCT 159568  
QY 60 eilePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGln 80  
DB 159567 CATCTTCATGCTAGGAATGCCCTCAATGGCGCGTCTTATGGCTCTCTGGTGCACAAAC 159508  
QY 80 rLysArgTrpSerCysAlaThrThrTyLeuValAsnLeuMetValAlaAspLeuLeu 100  
DB 159507 CAAGCCCTGGAGCTGTACCACTACCTACCTGATGAACCTGATGGTGGCCGACCTGTTA 159448  
QY 100 rValLeu---LeuProPheLeuIleIleThrTyLeuSerLeuAspArgTrpProPhe 119  
DB 159447 CATGCTGAGTCTGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 159388

QY 119 yGluLeuLeuCysLysLeuValHisPheLeuPheTyIleAsnLeuTyGlySerIleLe 139  
DB 159387 GGAGAGCGTGTGCAGCGTGGTTCCTGTCTTACACCACTCTACAGCAGCATCCT 159328  
QY 139 uLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLe 159  
DB 159327 GCTGCTGAGCTGCATCTCTATACCGCTTCTTAGGAGTGTGCCACCCCTGAGCTCACT 159268  
QY 159 uProTyArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValVa 179  
DB 159267 GTCCCTTCGGGACCCACCGCATACCTGGCAAGCAGCGCCACCTGGACCTGGTAGT 159208  
QY 179 lLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyIleAsnGlyGlnMetIl 199  
DB 159207 CGTCCAATTGCTACCGACTCTGGTCTTCTCCACACAGACTATGTAATGGCCACGTGAT 159148  
QY 199 eTrpTyArgMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyGlyIleValLe 219  
DB 159147 CTGCTACGACACAGCCAGCCAGAGAAATTTTCGATCATTTCTTCGCCCTACAGCATGGTCT 159088  
QY 219 uThrLeuSerGlyPhe--LeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGly 238  
DB 159087 GACACTATCTGGGTTGTCTTCCCTCTTGATCATTTCTCGGCTGCTACTCTCTGATGGT 159028  
QY 239 GlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGln 258  
DB 159027 CAGGACCTTGACAAATCCAGCAGGACCTCCCAAGCAGACGCGCCGCGTGCCAA 158968  
QY 259 ValHisProAspHisProThrGly 266  
DB 158967 GTCTATCCGACCATCTGCTGGT 158944  
RESULT 15  
AX244718/c  
LOCUS AX244718  
DEFINITION Sequence 47 from Patent WO0166750.  
ACCESSION AX244718  
VERSION AX244718.1 GI:15859597  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 681)  
AUTHORS Vogeli, G. and Wood, L. S.  
TITLE G protein-coupled receptors  
JOURNAL Patent: WO 0166750-A 47 13-SEP-2001;  
PHARMACIA & UPJOHN COMPANY (US)  
FEATURES  
Location/Qualifiers  
source 1..681  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 163 a 171 c 171 g 176 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,59e-57 Length: 681  
Score: 732.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 37.81% Indels: 0  
DB: 6 Gaps: 0  
US-10-023-586B-2 (1-360) x AX244718 (1-681)  
QY 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248  
DB 681 GGTCACTTTGGTGTGCTATCTACTGATGGTTCAGAGCGCTGATCAAGCCAGGAGAACCT 622  
QY 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268  
DB 621 CATGAGGACAGGCAACACAGCCGAGCGAGGTCCATCCGACCATCTCTACTGCTGTGG 562

Qy 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288  
|||||  
Db 561 CCTCTTCAACCTCTGTTTGTGGCCCTTCCATATCACTGGCTCCTTCTACCTCACCATCTG 502  
Qy 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308  
|||||  
Db 501 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTACAAGATATG 442  
Qy 309 GluAlaSerGlyGluCysGluLeuProGlnProSerProValLeuSerPheLysGly 328  
|||||  
Db 441 GAGGCCCTCTGTGTGAGTGTGAGCAGCTGCTCAACCCAGTCTGTACTTTCTTCAAGGGG 382  
Qy 329 GlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro 348  
|||||  
Db 381 GGCAAAATAGAGTCAGGCTCTCTCCAGAAACTGAGGCAGACAAAGTTGGGTGAGCATCCA 322  
Qy 349 AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
|||||  
Db 321 GCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 286

Search completed: May 27, 2003, 00:08:39  
Job time : 3925 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 21:29:43 ; Search time 304 Seconds  
(without alignments)  
2666.840 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILLPSRSGSRGAL.....QNKLGHPAGRRKRCPLNRS 360

oring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -dbv=xlh  
-O=/cgn2\_1/USPTO.spool/US10023586/runat.22052003\_114849\_28633/app\_query.fasta.1.519  
-DB=N\_Geneseq\_101002 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023586.ecgn.1.1.263.0runat.22052003\_114849\_28633 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6 -FGAPEXT=7  
-YGAPOB=10 -YGAPEXT=0.5 -DELOB=6 -DELEXT=7

Database :

N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1936	100.0	3143	24	AAS17746	Human genomic clon
2	1898	98.0	1063	24	AAS98135	Human DNA for pote
3	1798	92.9	1002	22	AH73516	Human G-protein-co
4	1785	92.2	1076	24	AAD29667	Human G-protein-co
5	1541	79.6	850	24	AAS17747	Human P2Y-like G-p
6	921	47.6	510	22	AAS50975	Human nGPCR12 codi
7	732	37.8	681	22	AAS30774	Human cDNA encodin
8	681	35.2	585	22	AAS30783	Human cDNA encodin
9	453	23.4	1163	22	AAD04981	Human purinergic r
10	445.5	23.0	1429	18	AAT74321	Human P2Y4 recepto
11	442	22.8	2025	22	AAD04980	Human purinergic r
12	420.5	21.7	984	17	AAT18367	Mouse pancreas G-p
13	416	21.5	1842	16	AAO88134	Human P20 receptor
14	415	21.4	2138	24	AAD32937	Mus musculus GPCR
15	415	21.4	2138	24	AAD32944	Mus musculus GPCR
16	413	21.3	984	17	AAT18368	Human placenta G-p
17	402	20.8	984	18	AAT70000	P20 purinergic rec
18	402	20.8	984	18	AAT747730	Human Placenta pur
19	370	19.1	551	22	AAS42837	Human G Protein-Co
20	328	16.9	6721	24	AAS18599	Purinergic recepto
21	327.5	16.9	6721	24	AAS18600	Purinergic recepto
22	294	15.2	450	21	AAA42873	Human secreted exp
23	292	15.1	234	22	AAS30782	Human cDNA encodin
24	281.5	14.5	1020	22	AAS51011	Human nGPCR54 codi
25	278.5	14.4	1011	24	AA43942	Human G-protein-co
26	278.5	14.4	1014	22	AAS07948	Human cDNA encodin
27	278.5	14.4	1014	24	ABO78847	Human G-protein co
28	278.5	14.4	1014	24	ABN85630	Human P2Y-like rec
29	278.5	14.4	1014	24	ABD34278	Human AXOR89 (G-pr
30	278.5	14.4	1014	24	ABK11381	Human DNA encoding
31	278.5	14.4	1014	24	AAK98323	Human purinergic-r
32	278.5	14.4	1014	24	AAD26370	Human G-protein co
33	278.5	14.4	1288	24	ABL56197	Human P2Y1-li enco
34	278.5	14.4	1729	22	AAS08362	Human cDNA encodin
35	278.5	14.4	1729	23	ABV24026	Human prostate exp
36	278.5	14.4	1729	23	ABV25767	Human prostate exp
37	278.5	14.4	1729	23	ABV29909	Human prostate exp
38	278.5	14.4	1729	23	ABV30024	Human P2Y1-li enco
39	278.5	14.4	5435	24	ABL56198	Human purinergic-r
40	278.5	14.4	9905	24	AAK98324	Human polynucleoti
41	278	14.4	1313	22	AAK52430	Human G-protein co
42	275	14.2	1301	17	AAT33905	Human G-protein co
43	275	14.2	1301	21	AAA70771	Human G-protein co
44	275	14.2	1301	21	AAZ60119	Human mu-opioid re
45	269	13.9	2162	19	AAV61993	

ALIGNMENTS

RESULT 1

AAS17746

ID AAS17746 standard; DNA; 3143 BP.

XX AAS17746;

XX 26-FEB-2002 (first entry)

DE Human genomic clone for P2Y-like G protein-coupled receptor.

XX Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;  
KW Chronic obstructive pulmonary disease; nervous system disease;  
KW Parkinson's disease; multiple sclerosis; dementia; stroke;  
KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;  
KW bacterial infection; fungal infection; protozoan infection;  
KW viral infection; pain; cancer; anorexia; bulimia; asthma;  
KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;  
KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;

psychotic disorder; neurological disorder; anxiety; schizophrenia;  
 manic depression; delirium; severe mental retardation; dyskinesia.

Homo sapiens.

Key Location/Qualifiers  
 CDS 520..2498  
 /\*tag= a  
 /product= "P2Y-like GPCR"

W0200185764-A2.

15-NOV-2001.

09-MAY-2001; 2001WO-EP05244.

11-MAY-2000; 2000US-203582P.

21-FEB-2001; 2001US-269857P.

(FARB ) BAYER AG.

Ramakrishnan S;

WPI: 2002-075242/10.

P-PSDB; AA011251.

New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders

Disclosure: Fig 1; 114pp; English.

The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR. In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, osteoporosis, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies, psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic depression, delirium, severe mental retardation or dyskinesias). The present sequence is a genomic clone encoding the P2Y-like GPCR of the invention.

Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 other;

#### Alignment Scores:

Pred. No.:	1-29e-174	Length:	3143
Score:	1936.00	Matches:	360
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-023-586B-2 (1-360) x AAS17746 (1-3143)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
 Db 436 ATGCTGTCCATTTTCTTCCAGGAGGAGCAGCAGGCGGCGTCTGAGCTCTG 495

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
 Db 436 CTCCTGGAGGAGCCCTCCGGGACATGGAGAAGTGGACATCAATACATCAGAGAACAA 555  
 Qy 41 GlyLeuCysGlnPheSerClnLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
 Db 556 GGTCTCTGCCAGTCTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 615  
 Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
 Db 616 ATCTTTATCTCTAGGCTGCCACTAAATGGCACTGCTGTGGCACTCTCTGGGGCCAAACC 675  
 Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
 Db 676 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCCGACCTGTTAT 735  
 Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120  
 Db 736 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 795  
 Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
 Db 796 CTGCTCTGCAAGCTGTGCACTTCTCTATATCAACCTTACGGCAGCATCTGCTG 855  
 Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
 Db 856 CTGACCTGCATCTCTGTGCACCACTTCTTAGTGTGTGCCACCACTGTGTTCGGCTGCC 915  
 Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeu 180  
 Db 916 TACCGACCCGAGGAGCTGCTGGCTGGCACCACTGCTGGCCCTGGTGTCTCTC 975  
 Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
 Db 976 CAGCTGTGCCACACTGGCTTCTCCACACGGACTACATCAATGGCCAGATGATCTGG 1035  
 Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
 Db 1036 TATGACATGACCAAGCAAGAGAAATTTTGTATCGCTTTTGGCTACGGCAGATGTTGACA 1095  
 Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
 Db 1096 TTGCTCTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1155  
 Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
 Db 1156 CCTGATCAAGCCAGAGAGAACCTCATGAGGACAGGCAACACACAGCCGAGCCAGTCCAT 1215  
 Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
 Db 1216 CCGGACCATCTTACTGTGTGTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1275  
 Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
 Db 1276 TCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1335  
 Qy 301 SerGlnCysGlyLeuLeuAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320  
 Db 1336 AGCCAGTGTGGCTTCAAGATATGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1395  
 Qy 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340  
 Db 1396 AGTCTCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTACAGCTCTCTCAGAAATGAGG 1455  
 Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
 Db 1456 CAGACCAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGAGGTGTGACAGATCT 1515  
 RESULT 2  
 AAS98135  
 ID AAS98135 standard; DNA; 1063 BP.  
 XX  
 AC AAS98135;

Alignment Scores:									
Pred. No.:	1.17e-171	Length:	1063						
Score:	1898.00	Matches:	354						
Percent Similarity:	99.72%	Conservative:	0						
Best Local Similarity:	99.72%	Mismatches:	0						
Query Match:	98.04%	Indels:	1						
DB:	24	Gaps:	0						
US-10-023-586B-2 (1-360) x AAS98135 (1-1063)									
Qy	7	ProSerArgGlySerArgSerGlySerArgArgGlyAlaLeuLeuLeuGluGlyAlaSer	26						
Db	1	CCTTCAGGGACACAGCGGAGCGCTGGAGCTCTGCTCTCGAGGAGGCTCC	60						
Qy	27	ArgAspMetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSer	46						
Db	61	CGGACATGGAGAGGTGGACATGATATACATCAGAGAACAGGTCTCTGCCAGTTCTCA	120						
Qy	47	GluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeu	66						
Db	121	GAGAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTTAGGGCTG	180						
Qy	67	ProLeuAsnGlyThrValLeuTyrPheHisSerTrpGlyGlnThrLysArgTrpSerCysAla	86						
Db	181	CCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCAACCAAGCGCTGGAGCTGTGC	240						
Qy	87	ThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeu	106						
Db	241	ACCACCTATCTGGTGAACCTGATGGTGGCCGACCTGCTTTATCTGCTATTGCCCTTCC	300						
Qy	107	IleIleThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuVal	126						
Db	301	ATCATCACTACTCCTAGATGACAGGTGGCCCTTCGGGAGCTGCTCTGCAAGCTGGTG	360						
Qy	127	HisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerVal	146						
Db	361	CACCTCTCTGTTCTATATCAACCTTTACGGCAGCATCCTGCTGTGACCTGCATCTCTGTG	420						
Qy	147	HisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArgTrpArgArgHis	166						
Db	421	CACCAGTCTCTAGGTGTGCCACCCACTGTTCGCTGCGCTACCGGACCCGACGAT	480						
Qy	167	AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu	186						
Db	481	GCCTGGCTGGGCACACACACACCTCGGCGCCTGGTGCTCTCCAGTGTCTGCCACACTG	540						
Qy	187	AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln	206						
Db	541	GCCTTCTCCCAACGCGACTACATCATATGGCCAGATGATCTGGTATGACATGACACCCAA	600						
Qy	207	GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSer	226						
Db	601	GAGAAATTTGATCGCCTTTTGCTACGGCATAGTCTGACATGTCTGGCTTTCTTTTCC	660						
Qy	227	LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGly	246						
Db	661	CTCCTTGCTCATTTTGGTGTGCTATTACCTACGTATTCACATGCTCCTCTTACCTCAC	720						
Qy	247	GluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGly	266						
Db	721	GAACCTCATGAGGACAGGCAACACAGCCCGAGCGTCCATCCGGACCATCTACTTGGT	780						
Qy	267	ValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHis	286						
Db	781	GTGTGGCCTCTTCACCTCTGTTTGTGGCCCTTCCATATCATTGCTCCTCTTACCTCAC	840						
Qy	287	HisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer-GlnCysGlyLeuGln	306						
Db	841	CATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTACA	900						
Qy	306	nAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPhe	326						
Db	901	AGATATGGAGCCCTCTGCTCAGTGTGAGCAGCTGCCTCAACCCAGTCTCTGCTACTTCTTT	960						

QY 326 elysGlyGlyLysAsnArgValArgLeuGlnLysLeuArgGlnAsnLysLeuGlyG1 346  
 Db 961 CAAGGGGGCAAAATAGAGTCAGGCTCTCCAGAAACTGAGGCAGACAAAGTTGGGTGA 1020  
 QY 346 uHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
 Db 1021 GCATCCAGCTGGAGGAAGAGATGCCAGGGTTGAACAGATCT 1063

## RESULT 3

AAH73516

ID AAH73516 standard; DNA; 1002 BP.

XX

AC AAH73516;

XX

DT 25-SEP-2001 (first entry)

XX

DE Human G protein-coupled receptor GPRV71 coding sequence.

XX

KW Human; guanosine triphosphate binding protein-coupled receptor;

KW G protein-coupled receptor; GPRV8; GPRV12; GPRV16; GPRV21; GPRV40;

KW GPRV47; GPRV71; GPRV72; cancer; liver cirrhosis;

KW Alzheimer's disease; cytostatic; hepatotropic; neurotropic;

KW neuroprotective; gene therapy; peptide therapy; ds.

Homo sapiens.

XX

PN WO200148188-A1.

XX

PD 05-JUL-2001.

XX

XX 28-DEC-2000; 2000WO-JP09408.

PF

XX 28-DEC-1999; 99JP-0375152.

PR

XX 31-MAR-2000; 2000JP-0101339.

PR

XX (HELI-) HELIX RES INST.

PA

PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;

PI Sugiyama T, Kishimoto T, Kanzaki K, Yasuda S, Inoue Y;

XX

DR WPI; 2001-425662/45.

DR

XX P-PSDB; AAG64125.

XX

PT New DNA encoding guanosine triphosphate binding protein coupled

PT receptors and their expression products for screening potential

PT anticancer and neurotropic drugs and in diagnosis of these diseases

XX

PS Example 1; Page 142-143; 170pp; Japanese.

XX

CC The invention relates to nine human guanosine triphosphate binding

CC protein (G protein)-coupled receptors designated GPRV8, GPRV12, GPRV16,

CC GPRV21, GPRV40, GPRV47, GPRV51, GPRV71 and GPRV72, and to the

CC genes encoding them. These genes and proteins and antibodies against

CC the protein are useful in the treatment, prevention, diagnosis and

CC investigation of diseases associated with G protein-coupled receptors,

CC including cancer, cirrhosis of the liver and Alzheimer's disease.

CC The present sequence is a G protein-coupled receptor coding sequence

XX

SQ Sequence 1002 BP; 207 A; 293 C; 249 G; 253 T; 0 other;

XX

## Alignment Scores:

Pred. No.:	3.85e-162	Length:	1002
Score:	1798.00	Matches:	331
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	92.87%	Indels:	0
DB:	22	Gaps:	0

US-10-023-586B-2 (1-360) x AAH73516 (1-1002)

QY

29 MetGluLysValAspMetAsnThrSerGlnGluGlyLeuCysGlnPheSerGluLys 48

Db	1	ATGGAGAAGGTGGACATGAATACATCACAGGAACAAAGGTCTCTGCAGTCTTCACAGAAAG	60
QY	49	TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu	68
Db	61	TACAAGCAAGTCTACTCTCCCTGGCCTACATGATCATCTTTATCTAGGGGTGCACATA	120
QY	69	AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrTrp	88
Db	121	AATGCACCTGTCTTGTGCACCTTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCAC	180
QY	89	TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuIleIle	108
Db	181	TATCTGGTGAACCTGATGGTGGCCGACCTGCTTTATGCTATTGCTCCCTTCCTCATCATC	240
QY	109	ThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe	128
Db	241	ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCATTC	300
QY	129	LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln	148
Db	301	CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTGACCTGCATCTCTGTGCCAC	360
QY	149	PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrp	168
Db	361	TTCTTAGGTGTGCCACCCACTGTGTCTGCTGCCCTACCGAGCCCGCAGGATGCCTGG	420
QY	169	LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe	188
Db	421	CTGGCACCAGCACCACCTGGGCCCTGGTGGTCTCCAGCTGCTGCCACACATGGCCTTC	480
QY	189	SerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsn	208
Db	481	TCCCACACGGAGTACATCAATGGCCAGATGATCTGGTATGACATGACCAAGCAAGAGA	540
QY	209	PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu	228
Db	541	TTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATGTTCTGGCTTCTTTCCCTCCTT	600
QY	229	GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro	248
Db	601	GGTCATTTTGGTGTCTATTCACTGATGCTCAGGAGCCTGATCAAGCCAGAGACACCT	660
QY	249	HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp	268
Db	661	CATGAGGACAGCAACACAGCCGAGCGTCCATCGGACCATCTCTACTGTTGTGTGG	720
QY	269	ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu	288
Db	721	CCTCTTCAACCCTCTGTTTGTGGCCCTTCCATATCACTCGCTCTTCTACCTCACCATCTG	780
QY	289	LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet	308
Db	781	CTTCTCTGCTTCTCAGGACTGCCAGCTCTTGTATGCCAGCAGCTGTGGCTACAAGATATG	840
QY	309	GluAlaSerGlyGluCysGlnGlnLeuProGlnProSerProValLeuSerPheLysGly	328
Db	841	GAGGCCTCTGGTGTAGTGTGAGCAGCTGCTCAACCCAGTCTCTGTACTTTCTTTCAAGGG	900
QY	329	GlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro	348
Db	901	GGCAAAATAGAGTCAGGCTCTTCCAGAACTGAGGCAGAAACAGTTGGTGAGCATCCA	960
QY	349	AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360	
Db	961	GCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 996	

## RESULT 4

AAD29667

ID AAD29667 standard; cDNA; 1076 BP.

XX

AC AAD29667;

XX

DT 17-MAY-2002 (first entry)  
DE Human G-protein coupled receptor (GCREC-1) CDNA.  
XX  
KW Human: G-protein coupled receptor; GCREC-1; cell proliferative disorder;  
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
KW transgenic animal; allergy; gene therapy; hepatotropic; anorectic;  
KW neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;  
KW virucide; gene; ss.  
XX Homo sapiens.  
H  
Key Location/Qualifiers  
FT CDS 57..1076  
FT /\*tag= a  
FT /product= "Human GCREC-1"  
XX  
PN WO200210387-A2.  
PD 07-FEB-2002.  
XX  
PF 25-JUL-2001; 2001WO-US23433.  
XX  
PR 27-JUL-2000; 2000US-221478P.  
PR 03-AUG-2000; 2000US-223268P.  
PR 21-AUG-2000; 2000US-227054P.  
PR 08-SEP-2000; 2000US-231121P.  
PR 13-SEP-2000; 2000US-232243P.  
PR 15-SEP-2000; 2000US-232691P.  
PR 22-SEP-2000; 2000US-235146P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NR;  
PI Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L;  
PI Graul RC, Warren BA, Lee EA, Ding L;  
XX  
WPI; 2002-188744/24.  
P-PSDB; AA018640.  
PT New human G-protein coupled receptor polypeptide for diagnosis,  
PT prevention and treatment of cell proliferative, neurological, and  
PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and  
PT metabolic disorders  
XX  
PS Claim 5; Page 137-138; 150pp; English.  
XX  
CC The invention relates to novel human G-protein coupled receptors  
CC (GCREC) and their encoding polynucleotides. GCREC is useful as an  
CC immunogen for preparing monoclonal and polyclonal antibodies. GCREC is  
CC useful for diagnosing, treating and preventing a cell proliferative  
CC disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder  
CC (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's  
CC disease, Parkinson's disease), a cardiovascular disorder (e.g.,  
CC atherosclerosis, hypertension, myocardial infarction), gastrointestinal  
CC inflammatory disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/  
CC allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder  
CC (e.g., diabetes, obesity, osteoporosis), and viral infections. GCREC is  
CC useful in a number of drug screening techniques, and to analyse the  
CC proteome of a tissue or cell type. GCREC is useful for creating knockin  
CC humanised animals or transgenic animals to model human diseases, in  
CC somatic or germline gene therapy, to generate a transcript image of a  
CC tissue or cell type, for detecting differences in the chromosomal  
CC location due to translocation, inversion, etc., among normal, carrier  
CC or affected individuals, and as hybridization probes for mapping  
CC naturally occurring genomic sequences. GCREC is useful in Southern or

CC northern analysis, dot blot or other membrane-based technologies, in PCR  
CC technologies, in dipstick, pin, microarray, and in microarrays utilising fluids or tissues from  
CC (ELISA)-like assays, and in microarrays utilising fluids or tissues from  
CC patients to detect altered GCREC expression. The present sequence is  
CC human GCREC-1 CDNA.  
XX  
SQ Sequence 1076 BP; 211 A; 318 C; 274 G; 273 T; 0 other;  
Alignment Scores:  
Pred. No.: 7 44e-161 Length: 1076  
Score: 1785.00 Matches: 337  
Percent Similarity: 99.12% Conservative: 1  
Best Local Similarity: 98.83% Mismatches: 1  
Query Match: 92.20% Indels: 2  
DB: 24 Gaps: 0  
US-10-023-586B-2 (1-360) x AAD29667 (1-1076)  
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
DB 57 ATGCTGTCCATTTCCTCCAGGGGAAGCAGAGGCGGAGCGCTCGTGAGCTCTG 116  
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
DB 117 CTCCTGGAGGGAGCCTCCCGGACATGGAGAGGTGGACATGATACATCACAGGAACA 176  
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
DB 177 GGTCTCTGCCAGTTCACAGAGAGTACAGAACTCTCCCTGGCTTCACAGTATC 236  
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
DB 237 ATCTTTATCTAGGGCTGCCACTAAATGGCACCTGTCTTGTGGCACCTCTCGGGCCAAACC 296  
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
DB 297 AAGCGTGGAGCTGTGCCACCATCTATCTGGTGAACCTGTATGGTGGCCGACCTGCTTTAT 356  
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPropheGlyGlu 120  
DB 357 GTGCTATTGGCCCTTCCTCATCATCACCTACTACTAGATCACAGGTGGCCCTTCGGGGGAG 416  
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
DB 417 CTGCTCTGCAAGCTGGTGACATCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTG 476  
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
DB 477 CTGACCTGCATCTCTGTGCACCATCTCTAGGTGTGTGGCACCCTCTGCTGCTGCTGCC 536  
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180  
DB 537 TACCGGACCGCAGGATGCTGGTGGGACACGACACCATCTGGGCGCTTGGTGGTCTC 596  
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
DB 597 CAGCTGTGCTCCACACTGGCTTCTCCACACGAGCTACATCAATGATGATGATGATGATG 656  
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
DB 657 TATGACATGACACGACCAAGAGAAATTTTGTATGCGGCTTTTTCCTACGCGATAGTTCTGACA 716  
QY 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240  
DB 717 TTGCTGGCTTTCTTCCCTCCCTCTTGGTGTGCTATTTGGTGTGCTATTCACCTGATGGTCAGG 776  
QY 240 luProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260  
DB 777 AGCCTGATCAAGCCAGAGGAGAACCTCATCAGGACAGGCAACACGCCCCGACGAGTCC 836  
QY 260 isProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280  
DB 837 ATCCGGACCATCTACTGTGTGGCTCTTCCACCTCTCTTGTGTGTGTGTGTGTGTGTGTGT 896







QY 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308  
 Db 781 CTTTCCTCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTTACAGATATG 840

QY 309 GluAlaSer 311  
 Db 841 GAGGCTCT 849

## RESULT 6

AAH50975

ID AAH50975 standard; DNA; 510 BP.

XX

AC AAH50975;

XX

28-AUG-2001 (first entry)

XX Human nGPCR12 coding sequence.

XX

KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;  
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
 KW cardiovascular disease; proliferative disorder; hormonal disorder;  
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
 KW attention deficit-hyperactivity disorder/attention deficit disorder;  
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
 KW neuroprotective; ds.

XX Homo sapiens.

OS

XX WO200136473-A2.

XX

25-MAY-2001.

XX

16-NOV-2000; 2000WO-US31581.

XX

16-NOV-1999; 9905-0165838.

PR 17-NOV-1999; 9905-0166071.

PR 19-NOV-1999; 9905-0166678.

PR 28-DEC-1999; 9905-0173396.

PR 22-FEB-2000; 2000US-0184129.

PR 28-FEB-2000; 2000US-0185421.

PR 02-FEB-2000; 2000US-0185554.

PR 02-MAR-2000; 2000US-0186530.

PR 03-MAR-2000; 2000US-0186811.

PR 09-MAR-2000; 2000US-0188114.

PR 17-MAR-2000; 2000US-0190310.

PR 21-MAR-2000; 2000US-0190800.

PR 20-APR-2000; 2000US-0198568.

PR 02-MAY-2000; 2000US-0201190.

PR 08-MAY-2000; 2000US-0203111.

PR 25-MAY-2000; 2000US-0207094.

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX

PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;

PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;

XX WPI; 2001-389826/41.

DR P-PSDB; AAG80935.

XX

New G protein-coupled receptor (nGPCR-x) and its encoding

PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -

XX Claim 4; Page 78; 261pp; English.

XX

The present invention relates to novel G protein-coupled receptors

CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,

CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present

CC sequence is the coding sequence for one such G protein-coupled receptor.

CC GPCRs are also known as seven transmembrane receptors and function in

CC signal transduction. The nGPCRx coding sequences are useful for

CC screening a human to diagnose a disorder affecting the brain or a genetic

CC

CC predisposition, specifically schizophrenia. nGPCRx are useful for  
 CC identifying compounds useful for treating schizophrenia. Detection of  
 CC nGPCRx in a sample is useful as a diagnostic tool for diseases or  
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,  
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular  
 CC diseases, proliferative disorders and hormonal disorders. Modulators of  
 CC nGPCRx activity have the utility for treating neurological disorders,  
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity  
 CC disorder/attention deficit disorder), and neuronal disorders such as  
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
 CC Additional disorders include inflammatory conditions (e.g. Crohn's  
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,  
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.  
 CC inflammatory bowel disease.

XX

SQ Sequence 510 BP; 92 A; 156 C; 126 G; 136 T; 0 other;

## Alignment Scores:

Pred. No.: 8,91e-79 Length: 510

Score: 921.00 Matches: 170

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 47.57% Indels: 0

DB: 22 Gaps: 0

US-10-023-586B-2 (1-360) x AAH50975 (1-510)

QY 83 TrpSerCysAlaThrThrTyrrLeuValAsnLeuMetValAlaAspLeuLeuTyrrValLeu 102

Db 1 TGGAGCTGTGCCACCACTATCTGGTGAACCTGTATGTGGCCGACCTGCTTTATGTGCTA 60

QY 103 LeuProPheLeuLeuLeuThrTyrrSerLeuAspArgTrpPheGlyGluLeuLeu 122

Db 61 TTGCCCTTCCTCATCATCCTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120

QY 123 CysLysLeuValHisPheLeuPheTyrrIleAsnLeuTyrrGlySerIleLeuLeuLeuThr 142

Db 121 TGCAAGCTGGTGCACCTTCCTGTTCTATATCAACCTTTACGGCAGCATCCTGCTGCTGACC 180

QY 143 CysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrrArg 162

Db 181 TGCATCTCTGTCACCACTTCCTAGGTGTGTGCCACCACTGTGCTGCTGCTGCTGCTGCTG 240

QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeu 182

Db 241 ACCGCGAGGCATGCTGGCTGGGCGACGACGACCACTTGGGCCCTTGGTGGTCTCCAGCTG 300

QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrrIleAsnGlyGlnMetIleTrpTyrrAsp 202

Db 301 CTGCCACACTGGGCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGGTATGAC 360

QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrrGlyIleValLeuThrLeuSer 222

Db 361 ATGACCAAGCAAGAGAAATTTTGATCGGCTTTTGGCCATGATGTTCTGACATGCTCT 420

QY 223 GlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnLeuProAsp 242

Db 421 GGCTTTCTTCCCTTCCCTTGGTCAATTTTGGTGTGCTATTCTACTGATGTCAGAGCCTGAT 480

QY 243 GlnAlaArgGlyGluProHisGluAspArg 252

Db 481 CAAGCCAGGAGAGAACTCATGAGGACAGG 510

RESULT 7

AAH50975/c

ID AAH50975 standard; cDNA; 681 BP.

XX

AC AAH50975;

XX

DT 04-DEC-2001 (first entry)

XX Human cDNA encoding G protein-coupled receptor nGPCR-75.

DE

XX

KW Human; G protein-coupled receptor; nGPR-x; ss; antiviral; analgesic;  
KW cytosolic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive;  
KW antiparkinsonian; nootropic; neuroprotective; antidepressant;  
KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;  
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;  
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;  
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;  
KW schizophrenia; migraine; major depression; anxiety; mental disorder;  
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.  
XX  
OS Homo sapiens.  
XX  
XX WO200166750-A2.  
XX  
XX 13-SEP-2001.  
XX  
XX 08-MAR-2001; 2001WO-US07322.  
XX  
XX 08-MAR-2000; 2000US-0187581.  
XX 08-MAR-2000; 2000US-0187582.  
XX 08-MAR-2000; 2000US-0187714.  
XX 08-MAR-2000; 2000US-0187715.  
XX 08-MAR-2000; 2000US-0187825.  
XX 08-MAR-2000; 2000US-0187828.  
XX 08-MAR-2000; 2000US-0187829.  
XX 08-MAR-2000; 2000US-0187830.  
XX 08-MAR-2000; 2000US-0187833.  
XX 08-MAR-2000; 2000US-0187874.  
XX 08-MAR-2000; 2000US-0187930.  
XX 08-MAR-2000; 2000US-0188049.  
XX 08-MAR-2000; 2000US-0187929.  
XX 08-MAR-2000; 2000US-0187928.  
XX  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
XX Vogeli G, Wood LS;  
XX  
XX WPI; 2001-536778/59.  
XX P-PSDB; AAU19205.  
XX

Isolated nucleic acid molecules encoding G protein-coupled receptors termed nGPR-x, useful in the treatment and diagnosis of viral infections, cancers and mental disorders (e.g. Parkinson's disease and schizophrenia) - ~~1~~.

Claim 4; Page 198; 336pp; English..

The invention relates to novel isolated nucleic acid molecules encoding G protein-coupled receptors termed nGPR-x. nGPR-x polynucleotides, and modulators may be used in the treatment of diseases and conditions such as infections, such as viral infections caused by HIV-1 (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, myocardial infarction, atherosclerosis), Parkinson's disease, and psychotic and neurological disorders, including schizophrenia, migraine, major depression, anxiety, mental disorder, manic depression, and dyskinesias, such as Huntington's disease or Tourette's Syndrome and many other diseases and syndromes listed in the specification.

nGPR-x polynucleotides and polypeptides, as well as nGPR-x modulators, may also be used in diagnostic assays for such diseases or conditions. The present sequence encodes a G protein-coupled receptor of the invention.

Alignment Scores:  
Pred. No.: 1.53e-60 Length: 681  
Score: 732.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 37.81% Indels: 0

Sequence 681 BP; 163 A; 171 C; 171 G; 176 T; 0 other;

DB: 22 Gaps: 0  
US-10-023-586B-2 (1-360) x AAS30774 (1-681)  
Qy 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248  
Db 681 GGTCACTTTGGTGTCTATTCTACCTGATGGTCAGGAGCCTGATCAAGCCAGAGGAGAACT 622  
Qy 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268  
Db 621 CATGAGGACAGCAACACAGCCGAGCCAGGTCATCCGGACCATCTTACTGTGTGTGG 562  
Qy 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288  
Db 561 CCTCTTCACTCTCTGTTTGTGCTTCCATATCATCTGCTCTTCTACCTACCATCTG 502  
Qy 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308  
Db 501 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGCACCCAGCTGTGGCTACAGATATG 442  
Qy 309 GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheLysGly 328  
Db 441 GAGGCTCTGCTGT 382  
Qy 329 GlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro 348  
Db 381 GGCAAAATAGAGT 322  
Qy 349 AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
Db 321 GCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 286  
RESULT 8  
AAS30783  
ID AAS30783 standard; cDNA; 585 BP.  
XX  
AC AAS30783;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Human cDNA encoding G protein-coupled receptor nGPR-84.  
XX  
KW Human; G protein-coupled receptor; nGPR-x; ss; antiviral; analgesic;  
KW cytosolic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive;  
KW antiparkinsonian; nootropic; neuroprotective; antidepressant;  
KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;  
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;  
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;  
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;  
KW schizophrenia; migraine; major depression; anxiety; mental disorder;  
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.  
XX  
OS Homo sapiens.  
XX  
XX WO200166750-A2.  
XX  
XX 13-SEP-2001.  
XX  
XX 08-MAR-2001; 2001WO-US07322.  
XX  
XX 08-MAR-2000; 2000US-0187581.  
XX 08-MAR-2000; 2000US-0187582.  
XX 08-MAR-2000; 2000US-0187714.  
XX 08-MAR-2000; 2000US-0187715.  
XX 08-MAR-2000; 2000US-0187825.  
XX 08-MAR-2000; 2000US-0187828.  
XX 08-MAR-2000; 2000US-0187829.  
XX 08-MAR-2000; 2000US-0187830.  
XX 08-MAR-2000; 2000US-0187833.  
XX 08-MAR-2000; 2000US-0187874.  
XX 08-MAR-2000; 2000US-0187930.  
XX 08-MAR-2000; 2000US-0188049.  
XX 08-MAR-2000; 2000US-0189294.



Qy	6	LeuPro---SerArgGlySerArgSerGlySerArgGlyAlaLeuLeuLeuGluGly	24
Db	160	CTTCCCATCTCCAGGGGGCCATGCGCCAGTACAGAGTCTCCCTGTTGAGATCCCTAGGC	219
Qy	25	AlaSerArgAspMetGluValAspMetAsnThrSerGlnGluGlnGlyLeuCysGln	44

```
Db 220 CTCAGC-----CCAGGTCCTGGCAGCAGTGCAGTGGAGCTGGAGCTGTGG 264
Qy 45 PheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeu 64
Db 265 TTTGATGAGATTCAAGTTTCCTGCTCCCTGCTGAGCTATGAGTGTCTTGTGGCTG 324
Qy 65 GlyLeuProLeuAsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSer 84
Db 325 GGCTTGGGCTTAAGGCCCAACCCATATGCTCTTCATCTTCGCGCTCCGACCCCTGGAT 384
Qy 85 CysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---Leu 103
Db 385 GCAACGCCCACTACATGTTTCCACTGGCATCTGCACACCTGTATGTATGCTGCTG 444
Qy 104 ProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys 123
Db 445 CCACCTCATCTACTATTATGAGCCCAACCACTGGCCCTTGGCACTGAGATCTGC 504
Qy 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys 143
Db 505 AAGTTCCTCGCTTCTTCTTCTATGTAACCTCTACTGCAGTGTCTTCTTCTCCACCTGC 564
Qy 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThr 163
Db 565 ATCAGCGTCGACCGCTACCTGGGCATCTGCCACCCACTTCCGGCACTACCTGGGCGCG 624
Qy 164 ArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuLeuLeu 183
Db 625 CCTCGCTCCGAGCCTTCTCGCTGGCAGTTTGGTGTGTGTAGCCGCTGCCCTCGTG 684
Qy 184 ProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMet 203
Db 685 CCCAACCTGTCTTGTGCACACCAAGAGGACCAAGGACCGCTGCTGCTGCCATGACACC 744
Qy 204 ThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGly 223
Db 745 ACTCGGCTGAAGAGTTTGACCA-CTATGTGCA-----CCTCAGCTCGCGCGT 791
Qy 224 -----PheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGly 238
Db 792 CATGGGCTGCTCTTGGCGCTGCCCTGCTGCTCACTCTTGTGTGTATGGACTCATGGC 851
Qy 239 GlnGluProAspGlnAlaArgGly-----GluProHisGluAsp 251
Db 852 TCGTCGCTGTATCAGCGCTTGGCAGGCTCTGCACAGCTGCTCTTCTCGCCT----- 902
Qy 252 ArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrpProLeuHis 271
Db 903 -----CCGCTCTCTCCGACCACTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 944
Qy 272 ProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeuLeu----- 289
Db 945 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 992
Qy 290 ---SerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 993 CAGGCTTTGGAGAGCTGACTCGCGAGTACTGACATTTGTCAACGTGCTCTATAAAGTGC 1052
Qy 309 GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheLysGly 328
Db 1053 TCGGCGCTGCGCAGTCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1112
RESULT 11
AAD04980
ID AAD04980 standard; DNA; 2025 BP.
XX
AC AAD04980;
XX
XX 17-JUL-2001 (first entry)
XX
XX Human purinergic receptor P2Y2 DNA.
XX
XX Human; purinergic receptor; P2Y2; antidepressant; vulnerary; hypotensive;
```

```
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;
KW irritable bowel disorder; reproductive system disorder; hypertension;
KW peripheral vascular disease; immune system disorder; chronic bronchitis;
KW premature ejaculation; asthma; neuromuscular disease; ds.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 246..1379
FT /tag= a
FT /product= "Human P2Y2 receptor protein"
FT primer_bind complement (246..263)
FT /tag= b
FT primer_bind /bound_moiety= "Sense primer"
FT /tag= c
FT primer_bind 1362..1379
FT /bound_moiety= "Antisense primer"
FT
FT US6214581-B1.
PN
XX 10-APR-2001.
XX
XX 13-NOV-1998; 98US-0191136.
XX
XX 16-JAN-1998; 98US-0071298.
PR 16-JAN-1998; 98US-0071669.
PR 16-JAN-1998; 98US-0008185.
PR 16-JAN-1998; 98US-0008526.
XX
PA (ABBO ) ABBOTT LAB.
XX
XX Lynch KJ, Burgard EC, Van Biesen T;
XX
XX WPI: 2001-315459/33.
DR P-PSDB; AAE01143.
XX
XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor
XX polypeptide useful for identifying potentially therapeutic compounds
XX that modulate or otherwise interact with P2X containing receptors -
XX
XX Example 14; Fig 11A; 53pp; English.
XX
XX The present sequence is a DNA encoding human wild type P2X4 receptor
XX protein. P2X receptors are ligand-gated ion channels while P2Y receptors
XX operate generally through a G-protein coupled system. P2X purinoreceptor
XX drugs are potential therapeutic agents in several disorders including
XX central nervous system or peripheral nervous system conditions, e.g.,
XX epilepsy, pain, depression, neurodegenerative disorders, disorders of the
XX skeletal muscle such as neuromuscular diseases, disorders of the
XX reproductive system, asthma, peripheral vascular disease, hypertension,
XX immune system disorders, irritable bowel disorder, premature ejaculation,
XX cystic fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the
XX activity of extracellular nucleotide triphosphates to regulate chloride
XX secretion in human airway epithelia.
XX
XX Sequence 2025 BP; 388 A; 617 C; 595 G; 425 T; 0 other;
```

## Alignment Scores:

Pred. No.:	3,69e-32	Length:	2025
Score:	442.00	Matches:	128
Percent Similarity:	50.91%	Conservative:	40
Best Local Similarity:	38.79%	Mismatches:	138
Query Match:	22.83%	Indels:	25
DB:	22	Gaps:	10

US-10-023-586B-2 (1-360) x AAD04980 (1-2025)

Qy	43	CysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe	62
Db	318	TGGCGCTTCAACGAGGACTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	377
Qy	63	IleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArg	82







XX	10-JUL-2000;	2000US-217058P.
PR	10-JUL-2000;	2000US-217179P.
PR	10-JUL-2000;	2000US-217223P.
PR	10-JUL-2000;	2000US-217253P.
PR	10-JUL-2000;	2000US-217255P.
PR	10-JUL-2000;	2000US-217256P.
PR	10-JUL-2000;	2000US-217257P.
PR	11-JUL-2000;	2000US-217347P.
PR	11-JUL-2000;	2000US-217629P.
PR	12-JUL-2000;	2000US-217537P.
PR	12-JUL-2000;	2000US-218069P.
PR	12-JUL-2000;	2000US-218074P.
PR	12-JUL-2000;	2000US-218358P.
PR	27-JUL-2000;	2000US-221483P.
PR	07-AUG-2000;	2000US-223120P.
PR	07-AUG-2000;	2000US-223122P.
PR	26-OCT-2000;	2000US-243958P.
PR	15-NOV-2000;	2000US-249408P.
PR	20-NOV-2000;	2000US-252299P.
PR	16-JAN-2001;	2001US-262113P.
PR	16-JAN-2001;	2001US-262205P.
XX	(DELT-)	DELTAGEN INC.
XX	PA	
XX	Allen KD,	Brennan TJ;
XX	PI	
XX	DR	
DR	WPI:	2002-164574/21.
DR	P-PSDB;	AAE20604.
XX	Novel non-human transgenic animal, especially transgenic mice useful for identifying an agent that modulates expression or function of target gene, comprises disruptions in target G protein-coupled receptor gene	
PT		
PT		
XX		
PS	Disclosure;	Fig 19; 101pp; English.
XX		
CC	The invention relates to a non-human transgenic animal having targetted G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene	
CC	5-HT-2B gene, chemokine receptor 9A gene, glucocorticoid-induced receptor	
CC	gene, orphan GPR10 (UHR-1) gene, orphan GPR14 gene, orphan GPR15 gene,	
CC	beta chemokine receptor (E01) gene, endothelial differentiation GPCR3	
CC	(EDG3) gene, ATP receptor P2U1 gene or adenosine 3 receptor gene. The	
CC	transgenic animal is useful for identifying an agent that modulates the	
CC	expression or function of the target gene, for identifying an agent that	
CC	ameliorates a phenotype associated with a disruption in the target gene.	
CC	A transgenic construct is useful for producing a transgenic animal,	
CC	preferably a transgenic mouse. The transgenic animal is useful for	
CC	testing the efficacy of proposed genetic and pharmacological therapies	
CC	for human genetic diseases, such as neurological, neuropsychological or	
CC	psychotic illnesses. The transgenic animal is also useful as models for	
CC	diseases, disorders or conditions associated with phenotypes relating to	
CC	a disruption in a target, and to identify pharmaceuticals, therapies,	
CC	drugs and interventions which may be effective in treating a disease or	
CC	other phenotypic characteristics of the animal. An agent which modulates	
CC	the expression of the target gene is useful as a therapeutic for treating	
CC	conditions associated with a disruption of the target gene. The present	
CC	sequence is mouse GPCR ATP-receptor P2U1 gene.	
XX		
SQ	Sequence 2138 BP; 423 A; 623 C; 606 G; 480 T; 6 other;	
	Alignment Scores:	
PR	Ad. No.:	1 51e-29
PR	Score:	415.00
PR	Percent Similarity:	49.69%
PR	Best Local Similarity:	36.50%
PR	Query Match:	21.44%
PR	DB:	24
	Length:	2138
	Matches:	119
	Conservative:	43
	Mismatches:	141
	Indels:	24
	Gaps:	6
US-10-023-586B-2 (1-360) x AAD32937 (1-2138)		
QY	43 CysGlnPheSerGluTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe 62	
	:    :      :            :    :    :    :	



```
Db 342 TGTGCTTCAACGAGGAGCTTCAAGTAGCTGTGTGCGCTGTCTATGCGGTGGTGC 401
Qy 63 ILeuGlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArg 82
Db 402 GTGCTGGGTGGCTGACGCGTGGCTCTCTATATCTTCCTATGCGCCCTCAAAACC 461
Qy 83 TrpSerCysAlaThrThrThrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102
Db 462 TGGAAACGCTCCACCACTACATGTTTCACCTGGCAGTTTCGGACTCTCTCTACGACGG 521
Qy 103 --LeuProPheLeuLeuLeuLeuThrThrSerLeuAspAspArgTrpPheGlyGluLeu 121
Db 522 TCCCTGCGCTGTGGTGTATTAATCTAGCCCGGGGTGACCACTGGCCATTAGCAGCGT 581
Qy 122 LeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeu 141
Db 582 CTCTGCAAGCTGGTGGCTTCTCTCTACACCACTCTACTGCACGATCCTCTTCCTC 641
Qy 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161
Db 642 ACCTGCATCAGCGTGCACCGGTGCTGGAGTCTCGCGCTCTGCACTCCCTGCGTGG 701
Qy 162 ArgThrArgArgHisAlaThrLeuGlyThrSerThrThrTrpAlaLeuValValLeuGln 181
Db 702 GCGCGGCGCGTATATGCGCGCGGTGGCTGCGGTGTGTGGTGTGGTGGCTGCG 761
Qy 182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyr 201
Db 762 CAGCACCGCTGCTCTACTTCGTCACCACTGCGGTGGGGAACCGGATCACTTGCAT 821
Qy 202 AspMetThrSerGlnGlnAsnPheAspArgPheAlaTyrGlyIleValLeuThrLeu 221
Db 822 GACACCTCGCGCGGAGAGCTCTTAGCATTCTGTGGCTTACAGCTCCGTCATGCTGGG 880
Qy 222 SerGlyPhe--LeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 881 TCTGCTTTTGTGCTGCGCTTTCCGTAATCCTGGTGTGTACGTGCTTATGGCCAGGCG 940
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 941 GCTGCTCAACCGGCTTATGGACCAAGAGGTCTCCCTCGGCGCAAGCGCAAGTCTGT 1000
Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db 1001 GCGCACCATTCCTTGTGCTGCGCGTCTTCGCGCTCTGCTTTCTGCTTTCACAGTCAC 1060
Qy 281 Ser-----LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAla 297
Db 1061 GCGCACCTCTACTACTCTCTTCG-----ATCACTTGACCTCAGCTGCCACAC 1108
Qy 298 LeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeu 317
Db 1109 CCTCAAGGCCATCAACATGGCATATAGATCAACCGCGCGCTGCGCCACGCAACAGTTG 1168
Qy 318 ProGlnPro-SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGln 337
Db 1169 -----TCTTGACCGGTACTCTACTTCTGCGGCGGACAGAGCTTGTCCGCTTGGCCG 1222
Qy 337 nLysLeuArg-----GlnAsnLysLeuGlyGln 346
Db 1223 AGATGCCAACCCACCGAGGAGCTACCCCGCCACAGGCTCGTCGCAAGCTGGGCGCT 1282
Qy 346 uHisProAlaGlyArg 351
Db 1283 GCACAGGCGCTAACAGA 1298
RESULT 15
ID AAD32944
AC AAD32944 standard; DNA: 2138 BP.
XX AC AAD32944;
XX 01-JUL-2002 (first entry)
```

```
XX Mus musculus GPCR ATP-receptor P2U1 gene targeting construct.
DE Mouse: G-protein coupled receptor; GPCR; transgenic animal; receptor;
KW genetic therapy; pharmacological; genetic disease; neuropsychological;
KW neurological; psychotic illness; nephrotropic; gynaecological;
KW psychostimulant; ATP-receptor P2U1; ds.
XX Mus musculus.
XX Key Location/Qualifiers
FT misc_feature 266..350
FT /tag= a
FT /note= "Sequence flanking Neo insert in targeting
FT misc_feature 351..636
FT /tag= b
FT /note= "Sequence deleted in targeting construct"
FT misc_feature 637..941
FT /tag= c
FT /note= "Sequence flanking Neo insert in targeting
FT construct"
XX WO200203793-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US21923.
XX 10-JUL-2000; 2000US-217058P.
XX 10-JUL-2000; 2000US-217179P.
XX 10-JUL-2000; 2000US-217233P.
XX 10-JUL-2000; 2000US-217253P.
XX 10-JUL-2000; 2000US-217255P.
XX 10-JUL-2000; 2000US-217256P.
XX 10-JUL-2000; 2000US-217257P.
XX 11-JUL-2000; 2000US-217347P.
XX 12-JUL-2000; 2000US-217537P.
XX 12-JUL-2000; 2000US-218069P.
XX 12-JUL-2000; 2000US-218074P.
XX 12-JUL-2000; 2000US-218358P.
XX 27-JUL-2000; 2000US-221483P.
XX 07-AUG-2000; 2000US-223120P.
XX 07-AUG-2000; 2000US-223122P.
XX 26-OCT-2000; 2000US-243958P.
XX 15-NOV-2000; 2000US-249408P.
XX 20-NOV-2000; 2000US-252299P.
XX 16-JAN-2001; 2001US-262113P.
XX 16-JAN-2001; 2001US-262205P.
XX (DELTA-) DELTAGEN INC.
XX Allen KD, Brennan TJ;
XX WPI; 2002-164574/21.
XX Novel non-human transgenic animal, especially transgenic mice useful
XX for identifying an agent that modulates expression or function of
XX target gene, comprises disruptions in target G protein-coupled receptor
XX gene.
XX Disclosure; Fig 20A; 101pp; English.
XX The invention relates to a non-human transgenic animal having targeted
XX G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene
XX 5-HT-2B gene, chemokine receptor 9A gene, glucocorticoid-induced receptor
XX gene, orphan GPR10 (UHR)-1 gene, orphan GPR14 gene, orphan GPR15 gene,
XX beta chemokine receptor (E01) gene, endothelial differentiation GPCR3
XX (EDG3) gene, ATP receptor P2U1 gene or adenosine 3 receptor gene. The
XX transgenic animal is useful for identifying an agent that modulates the
XX expression or function of the target gene, for identifying an agent that
XX ameliorates a phenotype associated with a disruption in the target gene.
```

A transgenic construct is useful for producing a transgenic animal, preferably a transgenic mouse. The transgenic animal is useful for testing the efficacy of proposed genetic and pharmacological therapies for human genetic diseases, such as neurological, neuropsychological or psychotic illnesses. The transgenic animal is also useful as models for diseases, disorders or conditions associated with phenotypes relating to a disruption in a target, and to identify pharmaceuticals, therapies, drugs and interventions which may be effective in treating a disease or other phenotypic characteristics of the animal. An agent which modulates the expression of the target gene is useful as a therapeutic for treating conditions associated with a disruption of the target gene. The present sequence is mouse GPCR ATP-receptor p2u1 gene targeting construct.

Sequence 2138 BP; 423 A; 623 C; 608 G; 480 T; 4 other;

Alignment Scores:	
Pred. No.:	1.51e-29
Score:	415.00
Percent Similarity:	49.69%
Best Local Similarity:	36.50%
Query Match:	21.44%
DB:	24
Length:	2138
Matches:	119
Conservative:	43
Mismatches:	141
Indels:	24
Gaps:	6

US-10-023-586B-2 (1-360) x AAD32944 (1-2138)

[illegible]

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 26, 2003, 22:54:23 ; Search time 58 seconds  
(without alignments)  
1903.511 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILLSRSGSRGRRGAL.....QNKLGHPAGRRKRCFLNRS 360

oring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5 ;

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.p2n.model -DEV=xlh

-O=/cgn2\_1/USPTO.spool/US10023586/runat\_22052003.114851.28697/app\_query.fasta\_1.519

-DB=Issued\_Patents\_NA\_QFMT-fastbp -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi

-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10023586.ecgn\_1.1.40 -runat\_22052003.114851.28697 -NCPU=6 -ICPU=3

-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120

-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	420.5	21.7	984	3	US-08-513-974B-41
2	420.5	21.7	1020	3	US-08-513-974B-370
3	420	21.7	1842	1	US-08-442-134A-1
4	420	21.7	1842	1	US-08-444-581B-1
5	420	21.7	1842	1	US-08-446-088A-1
6	413	21.3	984	3	US-08-513-974B-57
7	413	21.3	1023	3	US-08-513-974B-379
8	402	20.8	984	3	US-08-459-046-1
9	275	14.2	1301	2	US-08-467-948A-7
10	275	14.2	1301	3	US-08-467-947A-7
11	268	13.8	2160	4	US-08-188-275A-1
12	268	13.8	2162	4	US-09-351-198-1

13	268	13.8	2162	4	US-09-113-426-1	Sequence 1, Appl
14	267.5	13.8	1981	4	US-08-387-707-15	Sequence 15, Appl
15	267.5	13.8	1981	4	US-08-405-271A-15	Sequence 15, Appl
16	267	13.8	1610	3	US-08-889-108-7	Sequence 7, Appl
17	267	13.8	1610	5	PCT-US94-10358-7	Sequence 7, Appl
18	266.5	13.8	1901	1	US-08-153-848-43	Sequence 43, Appl
19	266.5	13.8	1901	3	US-09-299-843A-43	Sequence 43, Appl
20	266.5	13.8	1901	4	US-09-088-337B-43	Sequence 43, Appl
21	266.5	13.8	1901	5	PCT-US93-11153-43	Sequence 43, Appl
22	266.5	13.8	2453	5	PCT-US95-07180-1	Sequence 1, Appl
23	262.5	13.6	1205	1	US-08-417-103-13	Sequence 13, Appl
24	262.5	13.6	1634	1	US-07-816-283-1	Sequence 1, Appl
25	262.5	13.6	1634	1	US-08-417-103-1	Sequence 1, Appl
26	262.5	13.6	1910	2	US-09-009-438-1	Sequence 1, Appl
27	262.5	13.6	1910	3	US-09-207-493-1	Sequence 1, Appl
28	258.5	13.4	1640	3	US-08-781-250-1	Sequence 1, Appl
29	256	13.2	1265	1	US-07-816-283-3	Sequence 3, Appl
30	256	13.2	1265	1	US-08-417-103-3	Sequence 3, Appl
31	253.5	13.1	1102	2	US-08-742-440A-5	Sequence 5, Appl
32	253.5	13.1	1224	2	US-08-742-440A-4	Sequence 4, Appl
33	252	13.0	1255	1	US-08-097-938-3	Sequence 3, Appl
34	252	13.0	1255	1	US-08-476-000-3	Sequence 3, Appl
35	252	13.0	1255	1	US-08-472-840-3	Sequence 3, Appl
36	252	13.0	1255	2	US-08-476-976-3	Sequence 3, Appl
37	252	13.0	1255	3	US-08-474-410-3	Sequence 3, Appl
38	252	13.0	1255	4	US-08-486-673B-3	Sequence 3, Appl
39	250.5	12.9	1475	1	US-08-097-938-1	Sequence 1, Appl
40	250.5	12.9	1475	1	US-08-476-000-1	Sequence 1, Appl
41	250.5	12.9	1475	1	US-08-472-840-1	Sequence 1, Appl
42	250.5	12.9	1475	2	US-08-476-976-1	Sequence 1, Appl
43	250.5	12.9	1475	3	US-08-474-410-1	Sequence 1, Appl
44	250.5	12.9	1475	4	US-08-486-673B-1	Sequence 1, Appl
45	250.5	12.9	2732	1	US-08-476-000-60	Sequence 60, Appl

ALIGNMENTS

RESULT 1  
US-08-513-974B-41  
; Sequence 41, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiko  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995



CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/513,974B  
FILING DATE: 14-SEP-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01599  
FILING DATE: 10-AUG-1995  
APPLICATION DATA:  
APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995  
APPLICATION DATA:  
APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995  
APPLICATION DATA:  
APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1945  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 370:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1020 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..1011  
US-08-513-974B-370

Alignment Scores:  
Pred. No.: 6,1e-35  
Matches: 420/50  
Percent Similarity: 45.05%  
Best Local Similarity: 32.14%  
Query Watch: 21.72%  
DB: 3  
Length: 1020  
Matches: 117  
Conservative: 47  
Mismatch: 125  
Indels: 76  
Gaps: 9

US-10-023-586B-2 (1-360) x US-08-513-974B-370 (1-1020)  
QY 29 MetCluLysValAspMetAsnThrSerGlnGluGlnGlyLeu-----CysGln 44  
Db 28 ATGGAGCAG--GACAATGGCACCATCCAGGCTCCAGGCTCGCGCCACCACCACTGGCTG 84  
QY 45 PheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeu 64  
Db 85 TACCGTGAGGATTCACAGCGACTGCTCTAAACCCCGGTATACCTGGTGGTGGTGGTGC 144  
QY 65 GlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArgTrpSer 84  
Db 145 GGCCTGCCACTACATCTGCCTGCTATTCGCCAGATCTCGCATCCCGCCGACCTGACC 204  
QY 85 CysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---Leu 103  
Db 205 CGTTCCGCTGTGTACACCCTGAACCTGGCACTGGCGGACCTGATGATGCTGTTCACATA 264  
QY 104 ProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys 123  
Db 265 CCCCTACTTATCTATACTACGCCAGAGGGGACCACTGGCCCTTCGGAGACCTGGCCTGC 324  
QY 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys 143  
Db 325 CGCTTTGTAGCGCTTCTCTTATGCAATCTACATGGCAGCATCTGTTCTCCTCACCTGC 384  
QY 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCyssSerLeuProTyrArg--- 162  
Db 385 ATTAGCTTCAGCGCTACCTGGGCACTGCCACCCCTGGCTTCTCGCACACCGGTGGA 444  
QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuLeu 182  
Db 445 GGTGCCGCTGCTGCTGGTAGTGTGGAGTCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 504  
QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAsp 202  
Db 505 CTGCCACGCGAGTCTTTGTGTCACAGCATCCAGCGCAACCCACTGTGTGTGTGTGTGT 564  
QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer 222  
Db 565 CTGAGCCCAACCATCTCTACTCGCTACCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 624  
QY 223 GlyPheLeu---SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluPro 242  
Db 625 GGCTTCTTGTGCTGCCCTTCATAGCCTTACTGGCTTGTATTGTCTG----- 668  
QY 242 spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisPro 262  
Db 669 -----CATGGCCCGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 693  
QY 262 spHisProThr----- 265  
Db 694 GATGCCCAAGCAGGCTCTGTGGCCCAAGAGCGGCGCAGAGCGGCTCGTATGGCTGTG 753  
QY 266 -----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrTrpHisSerLeu 284  
Db 754 GTGGTGGCAGCTGCTTTGGCCATCAGCTT---CTGCTCTTCCACATCACCACAGCAGCC 810  
QY 284 euProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG 304  
Db 811 TACTTGGCTGTGCGCTCCACGCGGTGCTCTTGTGGCTGTGTGTGGAGACCTTGGCTGCT 870  
QY 304 LyLeuGlnAspMetGluAlaSerGlyGluCysGluGln----- 316  
Db 871 GCCTACAAAGGCACTCGGCTTCGCCAGGTGTCAACAGTGTTCGGACCCCATCTCTTTC 930  
QY 317 -----LeuProGlnProSerProValLeuSerPheLysGlyGlyLys 331  
Db 931 TACTTACACACAGAGAGTTCGGCGGC-ACCCCGCAGATCTCTACAGAGCTCACAG- 988  
QY 331 snArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAlaGly 351  
Db 989 -----CCAAGTGGCA 998









APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-513-974B-57

Alignment Scores:  
Pred. No.: 3,55e-34 Length: 984  
Score: 413.00 Matches: 126  
Percent Similarity: 48.61% Conservative: 31  
Best Local Similarity: 39.01% Mismatches: 146  
Query Match: 21.33% Indels: 22  
DB: 3 Gaps: 8

US-10-023-586b-2 (1-360) x US-08-513-974B-57 (1-984)

QY 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48  
Db 10 GACAATGGCACAGCCAGGCTCTGGGGCTTGCACCCACCACCTGTGTCTACCGGAGAAC 69  
QY 49 TyrLysGlnValTyrSerLeuAlaTyrSerIleIlePheLeuGlyLeuProLeu 68  
Db 70 TTCAAGCAACTGTGCTGCCACCTGTATTTCGGCGTGTGGCGGTGGCTGCCCTGCCGCTG 129  
QY 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
Db 130 AACATCTGTGTCTATACCCAGATCTGCACGCTCCCGCGGCGCTGACCCGCGCGCGGTG 189  
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 107  
Db 190 TACACCTAAACCTTGTCTGGGTGACCTGTATATGCTGCTCCCTGCCCTGCCCTGCTCATC 249  
QY 108 IleThrTrpSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127  
Db 250 TACAACCTATGCCCAAGGTGATCACTGGCCCTTTGGCGACCTTCGCTGCCGCTGGTCCGC 309  
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147  
Db 310 TTCCTCTCTATGCAACCTGCACGCGGAGCATCTCTTCCTCACTGCATCAGCTTCCAG 369  
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166  
Db 370 CGCTACCTGGGCATCTGCCACCCGCTGGCCCTTGGCACAAACCTGGGGCGCGCGGCT 429

QY 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186  
Db 430 GCCTGGCTAGTGTGTAAACCGTGTGGCTGGCCGTGACAAACCCAGTGCCTGCCACAGCC 489  
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206  
Db 490 ATCTTTCGACGACGATCCAGCATCCAGCGACTGTCTGTATGACCTCAGCCGCGCT 549  
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226  
Db 550 GCCCTGGCCACCCACTATATGCCCTATGGCATGGCTCTCACTGTCTGCGGCTTCTGCTG 609  
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246  
Db 610 CCCTTTCGCTGCCCTGGCTGGCTGCTACTGCTCTCTGGCGCTG---CCGCTTGGCGCCAG 666  
QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266  
Db 667 GATGGCCC-----GGCAGAGCCCTGTGGCCCGAGGCGGCGTGGCAGGCGCGCGCATG 720  
QY 266 lyValTrrp-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281  
Db 721 GCCG-TGGTGGTGGCTGCTGCTTTGCCATCAGCTT---CCTGCCCTTTTCATCATCACC 776  
QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301  
Db 777 GACAGCCTACCTGGCAGTGGCTCGACCGCGGCGTCCCTGCCTGTATTGGAGGCTT 836  
QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321  
Db 837 TGCAGCGCGCTTACAAAGGCAAGCGCGCTTTTGGCAGTGGCAACAGCGTGTGGACCCCAT 896  
QY 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuGlnLys 338  
Db 897 CCTCTTCTAC--TTACCCAGAGAGTTCGCGCGGCGGACCATGAGCTCCTACAGAAA 954  
QY 339 Leu 339  
Db 955 CTC 957

RESULT 7  
US-08-513-974B-379  
; Sequence 379, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 379:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1023 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 37..1020  
US-08-513-974B-379

Alignment Scores:  
Pred. No.: 3.78e-34 Length: 1023  
Score: 413.00 Matches: 126  
Percent Similarity: 48.61% Conservative: 31  
Query Local Similarity: 39.01% Mismatches: 146  
Query Match: 21.33% Indels: 22  
Gaps: 8

US-10-023-586B-2 (1-360) x US-08-513-974B-379 (1-1023)

Qy 33 AspMetAsnThrSerGlnGluClnGlyLeu-----CysGlnPheSerGluLys 48  
Db 46 GACAAATGACAGGAGGCTGTGGCTTGGCCACCCACACACCTGTGTCTACCGCGAGAAC 105  
Qy 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleLeuPheIleLeuPheLeuProLeu 68  
Db 106 TCAAGCAACTGCTCTGCCACCTGTATTCGGCGGTGTGGCGGCTGGCGGCTGGCGGCTG 165  
Qy 69 AsnGlyThrValLeuThrPheSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
Db 166 AACATCTGTGTCATTACCATCTGCACGTCCCGCGGCCCTGCACCGCAGCCCGCTG 225  
Qy 89 TyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 107

Db 226 TACACCCTAAACCTTGCTCTGCTGCTACCTGCTATATGCTGCTCCCTGCCCTGCTCATC 285  
Qy 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysValHis 127  
Db 286 TACAACATATGCCAAGGTGATCAGTGGCCCTTTGGCGACTTCCCTGCGCGCTGGCCGC 345  
Qy 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147  
Db 346 TTCCTCTTCTATGCCAACCTGCACGCGAGCATCCTTCTCCTCACCTGCATCAGTCCAG 405  
Qy 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166  
Db 406 CGCTACCTGGGCATCTGCCACCGCTGGCCCTTGGCAGCAAACTGGCGCGCGCGCT 465  
Qy 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuLeuLeuProThrLeu 186  
Db 466 GCCTGGCTAGTGTGTAACTGTGGTGGCGGTGCGGTGCAACCCAGTGCCTGCCACAGCC 525  
Qy 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206  
Db 526 ATCTTCGCTGCCACAGGATCCAGCGTAAACCGCATGCTGCTGCTATGACCTACCGCGCT 585  
Qy 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226  
Db 586 GGCCTGGCGCACCCACTATATGCCCTATGCGTATGCGTCTCCTGCTGCTGCTGCTG 645  
Qy 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246  
Db 646 CCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702  
Qy 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266  
Db 703 GATGGCCC-----GGCAGAGCGTCTGGCCAGGCGGTGGCAAGCGCGCGCATG 756  
Qy 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProThrHisSer 281  
Db 757 GCCG-TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 812  
Qy 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301  
Db 813 GACAGCTACCTGGCAGTGGGCTCGACGCGCGGCTGCCCTGCATGTTATGGAGGCTT 872  
Qy 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321  
Db 873 TGCAGCGCGCTACAAGGCGCAGCGCGCGCTTGGCAGTGCACAGCGTCTGGACCCCAT 932  
Qy 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuGlnLys 338  
Db 933 CCTCTTCTAC---TTCACCCAGAAAGTTCCCGCGCGCACATGAGCTCTACAGAA 990  
Qy 339 Leu 339  
Db 991 CTC 993

## RESULT 8

US-08-459-046-1  
Sequence 1, Application US/08459046  
Patent No. 6008039

## GENERAL INFORMATION:

APPLICANT: Coleman, Roger  
APPLICANT: Au-Young, Janice  
APPLICANT: Stuart, Susan G.  
TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3330 Hillview Avenue  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,046  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PF-0038 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY: Placenta  
CLONE: 179696  
US-08-459-046-1

```

Qy 207 GluAsnPheAspArgLeuPheAlaTyrglyIleValLeuThrLeuSerGlyPheLeu---S 220
Db 550 GCCCTGGCCACCCACTATATGCCCTATGGATGGCTCTCACTGTCATCGCTTCCTGCTG 609
Qy 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluPProAspGlnAlaArgG 246
Db 610 CCCCTTTCCTGCCCTGCTGCTCTACTGCTCTGCTGGCCTG---CCGCTGTGCCGCCAG 666
Qy 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
Db 667 GATGCCCC-----GGCAGAGCCTGTGGCCACGAGCGGGCTGGCAAGCGGCCGCCGATG 720
Qy 266 ly-----ValTrpProLeuHisProLeuPheCysAlaLeuProTyH 280
Db 721 GCCCTGGTGGCTGCTGCTTTTGGCATCAGCTTCTGCTCTT-----TCACATC 771
Qy 280 isserLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
Db 772 ACCAAGACACCTACCTGGCAGTGCCTCGACCGCGGCTGCCCTGTCACATGTATTGGAG 831
Qy 300 lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
Db 832 GCCTTTCACCGGCTACAAGACGCGGCCCTTTGCCAGTGCCCAACAGCGTGTGGAC 891
Qy 320 roserProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuG 337
Db 892 CCCATCCTCTTCTAC--TTCACCCACAGAAGAAGTTCCGCGGCGGACCATGAGCTCCTAC 949
Qy 337 lnlysLeu 339
Db 950 AGAAATC 957

RESULT 9
US-08-467-948A-7
: Sequence 7, Application US/08467948A
: Patent No. 5998164
: GENERAL INFORMATION:
: APPLICANT: LI, YI
: APPLICANT: CAO, LIANG
: APPLICANT: NI, JIAN
: APPLICANT: GENTZ, REINER
: APPLICANT: BULT, CAROL J.
: APPLICANT: SUTTON III, GRANGER G.
: APPLICANT: ROSEN, CRAIG A.
: TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
: TITLE OF INVENTION: Coupled Receptor GPR2
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
: STREET: 1100 NEW YORK AVE., NW, SUITE 600
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY DISK
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,948A
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04079
: FILING DATE: 30-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: STEFFE, ERIC K.
: REGISTRATION NUMBER: 36,688
: REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-2600
: TELEFAX: 202-371-2540

```

## ; INFORMATION FOR SEQ ID NO: 7:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: .. both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 161..1192
US-08-467-948A-7
```

## Alignment Scores:

```
Pred. No.: 1.92e-19 Length: 1301
Score: 275.00 Matches: 88
Percent Similarity: 47.72% Conservative: 48
Best Local Similarity: 30.88% Mismatches: 106
Query Match: 14.20% Indels: 44
DB: 2 Gaps: 10
```

US-10-023-586B-2 (1-360) x US-08-467-948A-7 (1-1301)

```
QY 23 GluGlyAlaSer-----ArgAspMetGluLysValAspMetAsnThrSer 37
Db 122 GAAGTGGCTTCAAACTGAAATGACGCTGCTTTACGATGGTAAGCGTTACAGCTCC 181
Db 38 GlnGluGlnGlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAla 57
Db 182 CAC-----TGCTTCTATAATGACTCTTTAAGTACACTTGTATGGTGCATG 229
QY 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrp 77
Db 230 TTCAGCATGGTGTGTGGTGTAAATCAATGTGTGCCATATACATATTTTCATC 289
QY 78 GlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97
Db 290 TCGGCTCCTCAAGTCGGAATGAACACTCACTTACATGATTAACITGGCAATGTCAGAC 349
QY 98 LeuLeuValLeu-----LeuProPheLeuIleIleThrTyrSerLeuAspArgTrp 116
Db 350 TTGCTTTTGTGTTTACTTTTACCTTCAGGATTTT---TACTTCAACACGGAATGG 406
QY 117 ProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136
Db 407 CCATTGGAGATTTACTTTGTAAGATTTCTGTGCTGCTTTTATACCAACATGTACGGA 466
QY 137 SerIleLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156
Db 467 ACATCTCTGTTCTTAACCTGTATAGTAGATCGATTTCTGCAATGCTTACCCATTT 526
QY 157 CysSerLeuProTyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAla 176
Db 527 AAGTCAAGACTCTAAGAACCAAAAGAAATGCAAGATTTGTGCACTGGCGTGGTTA 586
QY 177 LeuValValLeuGlnLeuLeuProThr-----LeuAlaPheSerHisThrAspTyrIle 194
Db 587 ACTGTGATCGGAGAGTGCACCCCGCGTTTGTGTGCTACTACCCACTCTCAGGGTAAC 646
QY 195 AsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsnPheAsp-----210
Db 647 ATGCCTCAGAGCTGCTTT-----GAAATTTTCCAGAACCCACATCG 691
QY 211 ArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuGlyHis 230
Db 692 AAAACATATCTCTCAAGGATTTGTAAT-----TTCATCGAAATAGTGGGATTT 739
QY 231 Phe-----GlyValLeuPheThrAspGlyGln 239
Db 740 TTTATCTCTTAATTTTAATGTAACCTTGTCTAGTATGGTGTCTAAACACTTTAACCAA 799
QY 240 GluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnVal 259
Db 800 CTGTGTACATTAGTAGAGCAAAATAAACAAACTAAGCTTTTAAAAATGATTTTGTGTA 859
```

```
QY 260 HisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyr 279
Db 860 CATTG-GATCAT-----ATTCTGTTCTGTTTGTCTTCCCTTACAATAT 900
QY 280 HisSerLeuLeuLeu 284
Db 901 CAATCTTATTTTATA 915
```

## RESULT 10

```
US-08-467-947A-7
; Sequence 7, Application US/08467947A
; Patent No. 6090575
; GENERAL INFORMATION:
; APPLICANT: LI, VI
; APPLICANT: CAO, LIANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; TITLE OF INVENTION: Coupled Receptor GPRI
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,947A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 7:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 161..1192
US-08-467-947A-7
```

## Alignment Scores:

```
Pred. No.: 1.92e-19 Length: 1301
Score: 275.00 Matches: 88
Percent Similarity: 47.72% Conservative: 48
Best Local Similarity: 30.88% Mismatches: 106
Query Match: 14.20% Indels: 44
DB: 3 Gaps: 10
```

US-10-023-586B-2 (1-360) x US-08-467-947A-7 (1-1301)

QY 23 GluGlyAlaSer-----ArgAspMetGluLysValAspMetAsnThrSer 37

```
Db 122 GAAAGTCTTCCAAACTGAAATTTGGACGTGCCTTTACGATGTTAGCGTTAAACAGCTCC 181
Qy 38 GlnGluGlnGlyLeuCysGlnPheSerGlyLysTyrLysGlnValTyrLeuSerLeuAla 57
Db 182 CAC-----TGCTCTATAATGACTCCTTTAAGTACACTTGTGTATGGGTGCGATG 229
Qy 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrp 77
Db 230 TTCAGCATGCTGTGTGCTGGGTTAATATCCAAATTTGCTGTCATATACATTTTCATC 289
Qy 78 GlyClnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97
Db 290 TGGCTCCTCAAGTCGGAATGAACTACAACTACATGATTAATCTGGCAATGTCAGAC 349
Qy 98 LeuLeuTyrValLeu---LeuProPheIleIleThrTyrSerLeuAspAspArgTrp 116
Db 350 TTGCTTTTGTGTTTACTTTACCTTCAGGATTTT---TACTTCACAAACGGAATGG 406
Qy 117 ProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136
Db 407 CCATTTGGAGATTTACTTTTGAAGATTTCTGTGATGCTGTTTATACCAACATGTACGGA 466
Qy 137 SerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156
Db 467 AGCATTTCTGTTTAACTGTTAGTGTAGATCGATTTCTGCGAATTTGTCTACCCATTT 526
Qy 157 CysSerLeuProTyrArgTrpArgHisAlaTrpLeuGlyThrSerThrThrTrpAla 176
Db 527 AAGTCAAGACTCTAAGAACCAAAAGAAAGTGTGTTGTCACCTGGGTGTGTTA 586
Qy 177 LeuValValLeuGlnLeuLeuProThr-----LeuAlaPheSerHisThrAspTyrIle 194
Db 587 ACTGTGATCGGAGGAAGTGCACCGCGCTTTTGTGTCAGTCTACCCACTCTCAGGGTAAC 646
Qy 195 AsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGluAsnPheAsp-----210
Db 647 AATGCCCTCAGAGCGCTGCTTT-----GAAATTTTCCAGAGCCACATGG 691
Qy 211 ArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuGlyHis 230
Db 692 AAAACATATCTCTCAAGATTTGTAAT-----TTCATCGAAATAGTGGGATTT 739
Qy 231 Phe-----GlyValLeuPheThrAspGlyGln 239
Db 740 TTTATTCTCTAAATTTTAAATGTAATTTGTTCTAGTATGTTGCTTAAACCTTTAACCAAA 799
Qy 240 GluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnVal 259
Db 800 CCTGTTACATTAGTAGAACCAAAATAACAAACTAAGGTTTAAATAAGATTTTGTGA 859
Qy 260 HisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyr 279
Db 860 CATTT-GATCAT-----ATTCTGTTTCTGTTTGTCTTACATAT 900
Qy 280 HisSerLeuLeuLeu 284
Db 901 CAATCTTATTTTATA 915
RESULT 11
US-08-188-275A-1
; Sequence 1, Application US/08188275A
; Patent No. 6258556
; GENERAL INFORMATION:
; APPLICANT: Wang, Jia-Bei
; APPLICANT: Uhl, George R.
; APPLICANT: Johnson, Peter S.
; APPLICANT: Persico, Antonio
; TITLE OF INVENTION: cDNA and Genomic Clones Encoding Human
; TITLE OF INVENTION: Mu Opiate Receptor and the Purified Gene Product
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
```

```
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,275A
FILING DATE: 28-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1173-449P
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2160 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: -
LOCATION: 1..2160
OTHER INFORMATION: /label= cDNA
OTHER INFORMATION: /note= "cDNA encoding human mu opiate receptor"
US-08-188-275A-1
Alignment Scores:
Pred. No.: 2,34e-18 Length: 2160
Score: 268.00 Matches: 95
Percent Similarity: 46.71% Conservative: 47
Best Local Similarity: 31.25% Mismatches: 128
Query Match: 13.84% Indels: 35
DB: 4 Gaps: 12
US-10-023-586B-2 (1-360) x US-08-188-275A-1 (1-2160)
Qy 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrp 77
Db 441 TACTCCATCGTGTGGTGGGCTCTTCGGAACCTTCCTGGTATGATGTATGTGATGTC 500
Qy 78 GlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97
Db 501 AGATACACCAAGATGAGACTGCCACCAACATCTACATTTTCAACCTTGCTGTCGCAGAT 560
Qy 98 LeuLeuTyrVal---LeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrp 116
Db 561 GCCTTAGCCACCACTACCTGCCCTTCCAGAGTGTGAATTAAC---CTAATGGGAACATGG 617
Qy 117 ProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136
Db 618 CCATTTGGAACCACTCTTTGCAAGATAGTAGTCCATAGATTAATATATATAGTTGTCACC 677
Qy 137 SerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156
Db 678 AGCATATTCACCTCTGCACCATGAGTCTTGATCGATCATCATTCAGCTGTCACCCCTGTC 737
Qy 157 CysSerLeuProTyrArgTrpArgHisAlaTrpLeuGlyThrSerThrThrTrpAla 176
Db 738 AAGGCTTAGATTTTCCGCTACTCCCGAAATGCCAAATATCAATGCTGTGCAACTGGATC 797
Qy 177 LeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGly 196
Db 798 CTCTCTTCAGCCATTTGGTCTTCTCTTAATGTTTCATGCTACACAAATAACAGGCAAGT 857
```

```
Qy 197 GlnMet-----iletpyrAspMetThrSerGlnGlu 207
Db 858 TCCATAGATTGTACACTAACATTTCTCTCATCCAACTGGTAC-----TGGGAA 905

Qy 208 AsnPheAspArgLeuPheAlaTyrGlylleValLeuThrLeuSerGlyPheLeuSerLeu 227
Db 906 AACCTCGTGAAGATC-----TGTGTTTTCATCTTCCGCTTCATTATGCCAGTG 953

Qy 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247
Db 954 CTCAT-CATTACCGTGTGCTAGGACTGATGATCTTGGCCCTCAAGAGTGT-----1003

Qy 248 ProHis-----GluAspArgGlnHisSerProSerGlnValHisProAspHis 263
Db 1004 CGCATGCTCTCTGCTGCCAAAGAAAGACAGGAATCTTCGAAGGATCACCAGGATGT 1063

Qy 264 ProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeu 283
Db 1064 GCTGTGTGTGGTGGTGTGTTTCATCTCTG-----CTGGACTCCCATTTCA---CAT 1111

Qy 284 LeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCys 303
Db 1112 TTACGTGCATCTTAAGCCTTGGTTACAT-----CCAGAACTAGTTCACAGCTGT 1165

Qy 304 GlyLeuGlnAspMetGluAlaSerGlyGlyCysGluGlnLeuProGlnProSerProVal 323
Db 1166 TTCTTGGCACTTCTGCATTTGCTAGTGTACAAACAGCTGCCCTCAACCCAGTCTCT---1222

Qy 324 LeuSerPheLeuGlyGlyAsnArgValArgLeuGlnLeuLysLeuArgGlnAsnLys 343
Db 1223 TTATGCCATTTCTGGATGAAACTTCAAACG---ATGCTTCAGAGAGTCTGTATCCCAAC 1279

Qy 344 LeuGlyGluHis 347
Db 1280 CTCTTCCAACAT 1291

RESULT 12
US-09-351-198-1
; Sequence 1, Application US/09351198
; Patent No. 6335168
; GENERAL INFORMATION:
; APPLICANT: Kreek, Mary J
; APPLICANT: Laforge, Karl S
; APPLICANT: Yu, Lei
; TITLE OF INVENTION: METHODS OF THE HUMAN MU OPIOID RECEPTOR, DIAGNOSTIC
; TITLE OF INVENTION: METHODS OF USING SAID ALLELES, AND METHODS OF TREATMENT
; TITLE OF INVENTION: BASED THEREON
; FILE REFERENCE: 600-1-226N
; CURRENT APPLICATION NUMBER: US/09/351,198
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/092,402
; EARLIER FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2162
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2063)
; OTHER INFORMATION: No. 6335168feature for this position in GeneBank.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2091)
; OTHER INFORMATION: No. 6335168feature for this position in GeneBank.
; US-09-351-198-1

Alignment Scores:
Pred. No.: 2,35e-18 Length: 2162
Score: 268.00 Matches: 95
```

```
Percent Similarity: 46.71% Conservative: 47
Best Local Similarity: 31.25% Mismatches: 128
Query Match: 13.84% Indels: 35
DB: 4 Gaps: 12

US-10-023-586B-2 (1-360) x US-09-351-198-1 (1-2162)

Qy 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrp 77
Db 441 TACTCCATCGTGTGGTGGGCTCTTCGGAACATTCCTGGTCATGATGTATGATGTC 500

Qy 78 GlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97
Db 501 AGATACACCAAGATGAAGACTGCCACCAACATCTACATTTTCAACCTTGCTCTGCAGAT 560

Qy 98 LeuLeuTyrVal---LeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrp 116
Db 561 GCCTTAGCCACCATGACCTGCCCTTCCAGAGTGTGAATTAC---CTAATGGGAACATGG 617

Qy 117 ProPheGlyGluLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136
Db 618 CCATTTGGAACCATCTTTGCAAGATAGTGTCTCCATAGATTACTATATAACATGTTCAAC 677

Qy 137 SerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156
Db 678 AGCATATTACCCCTGTCACCATGATGTTGATCATATCATTTGCACCTCTGCCACCTGTC 737

Qy 157 CysSerLeuProTyrArgThrArgHisAlaTyrLeuGlyThrSerThrThrTrpAla 176
Db 738 AAGGCTTAGATTTCCGTACTCCCGAATGCCAAATATCAATGTCGCAACTGGATC 797

Qy 177 LeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGly 196
Db 798 CTCTCTCAGCCATTTGGTCTTCTCTGTAATGTTCACTGGCTACAAACAAATACAGGCAAGT 857

Qy 197 GlnMet-----iletpyrAspMetThrSerGlnGlu 207
Db 858 TCCATAGATTGTACACTAACATTTCTCTCATCCAACTGGTAC-----TGGGAA 905

Qy 208 AsnPheAspArgLeuPheAlaTyrGlylleValLeuThrLeuSerGlyPheLeuSerLeu 227
Db 906 AACCTCGTGAAGATC-----TGTGTTTTCATCTTCCGCTTCATTATGCCAGTG 953

Qy 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247
Db 954 CTCAT-CATTACCGTGTGCTAGGACTGATGATCTTGGCCCTCAAGAGTGT-----1003

Qy 248 ProHis-----GluAspArgGlnHisSerProSerGlnValHisProAspHis 263
Db 1004 CGCATGCTCTCTGCTGCCAAAGAAAGACAGGAATCTTCGAAGGATCACCAGGATGT 1063

Qy 264 ProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeu 283
Db 1064 GCTGTGTGTGGTGGTGTGTTTCATCTCTG-----CTGGACTCCCATTTCA---CAT 1111

Qy 284 LeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCys 303
Db 1112 TTACGTGCATCTTAAGCCTTGGTTACAT-----CCAGAACTAGTTCACAGCTGT 1165

Qy 304 GlyLeuGlnAspMetGluAlaSerGlyGlyCysGluGlnLeuProGlnProSerProVal 323
Db 1166 TTCTTGGCACTTCTGCATTTGCTAGTGTACAAACAGCTGCCCTCAACCCAGTCTCT---1222

Qy 324 LeuSerPheLeuGlyGlyAsnArgValArgLeuGlnLeuLysLeuArgGlnAsnLys 343
Db 1223 TTATGCCATTTCTGGATGAAACTTCAAACG---ATGCTTCAGAGAGTCTGTATCCCAAC 1279

Qy 344 LeuGlyGluHis 347
Db 1280 CTCTTCCAACAT 1291

RESULT 13
US-09-113-426-1
```





Percent Similarity: 42.28% Conservative: 49  
 Best Local Similarity: 29.00% Mismatches: 151  
 Query Match: 13.82% Indels: 63  
 DB: 4 Gaps: 14

US-10-023-586B-2 (1-360) x US-08-387-707-15 (1-1981)

QY 24 GlyAlaSerArgAspMetGluLysValaspMetAsnThrSerGlnGluGlnGly----- 41  
 DB 334 GGCTCTGGCTCAACTGTGCTCCACGTTGATGGAACACAGTCCGACCCATCGGTCTCTAAC 333  
 QY 42 -----LeuCysGlnPheSer-----GluLysTyr 49  
 DB 394 CGACGGGCTTGGCGGGAACGACACGCTGTGCCCTCAGACCGGAGCCCTCCATGGTC 453  
 QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGluProLeuAsn 69  
 DB 454 ACAGCCATCACCATGCGCCCTCTATTCTGCTGTGTGAGTGGCCCTCTTTGGAAAC 513  
 QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrTyr 89  
 DB 514 TTCTCTGTCATGTATGTGTTGTAAGATATACCAAAATGAAGACTGCCACCAACATCTAC 573  
 QY 90 LeuValAsnLeuMetValAlaAspLeuTyrVal---LeuLeuProPheLeuIle 108  
 DB 574 ATTTTCAACCTGCTCTGCAGATGCTTAGCCACTAGCAGCGCTGCCCTTTCAGAGTGT 633  
 QY 109 ThrTyrSerLeuAspArgTrpProPheGlyGluLeuLysLysLeuValHisPhe 128  
 DB 634 AACTAC---CTGATGGGAAGCTGGCCCTTGGAAACATCTCTGCAAGATCGTATCTCA 690  
 QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
 DB 691 ATAGACTACTACAACATGTGTACCAGTATCTTCAACCCTCTGCACCATGAGTGTAGACCG 750  
 QY 149 PheLeuGlyValCysHisProLysSerLeuProTyrArgThrArgHisAlaTrp 168  
 DB 751 TACATTGGCTGTGCCACCGGTCAGGCGCTGGATTCTCGTACCCCGCAAGATGCCAA 810  
 QY 169 LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188  
 DB 811 ATTGTCAATGTCTGCAACTGGATCTCTCTGCGCATTTGGTCTGCCGTATGTTCATG 870  
 QY 189 SerHisThrAspTyrIleAsnGlyGlnMet-----Ile 199  
 DB 871 GCAACACACAAAATACAGCAGGCGTCCATAGATTGACCCCTCAGTCTCTCATCCACA 930  
 QY 200 TrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeu 219  
 DB 931 TGGTAC-----TGGGGAACCTGCTCAAAATCTGTCTCTTCATCTTCGCTTC 978  
 QY 220 ThrLeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAsp----- 237  
 DB 979 ATCATGCGCGGC-----CTCAT-CATCACTGTGTGTTAGTACTGATGATCTT 1025  
 QY 238 ---GlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerPro 256  
 DB 1026 ACAGCTCAAGAGTGTCCGATGCTGTGCGGCTCCAAAGAAAGAGCAGGAA-----CCT 1079  
 QY 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLysPheCysAla 276  
 DB 1080 GCGCAG-----GATCACCAGGATGTGTGGTGGTGGTGTATTATTATGCTGTG 1130  
 QY 277 LeuProTyrHisSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuPro 296  
 DB 1131 CTGGACCCCATCCATCTATGTCTATCATCAAGACACTGATCAGCATTCAGAAACCCAC 1190  
 QY 297 AlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGln 316  
 DB 1191 TTTCCAGAC-----TGTTCCTGGCACTTCTCATTTGCTTGGGTATACAAACAG 1241  
 QY 317 LeuProGlnProSerProValLeuSerPheLysGlyGlyLys----- 330  
 DB 1242 CTGCCTGAACCCAGTTCT---TTATGGTTCTCTGGATGATGAACCTCAACAGATGTTTAG 1298  
 QY 331 -----AsnArgValArgLeuLeu-GlnLysLeuArgG1 341  
 DB 1299 AGAGTTCTGCATCCCAACTTCTCCACATCGAACACGAAACTCTGCTCGATCCGTCA 1358  
 QY 341 nAsnLysLeuGlyGluHisProAla 349  
 DB 1359 AAAC---ACTAGGGAACACCCCTCC 1380

RESULT 15  
 US-08-405-271A-15  
 ; Sequence 15, Application US/08405271A  
 ; Patent No. 6432652  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EVANS, CHRISTOPHER J.  
 ; APPLICANT: KEITH, DUANE E.  
 ; TITLE OF INVENTION: OPIOID RECEPTOR GENES  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORRISON & FOERSTER  
 ; STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500  
 ; CITY: WASHINGTON  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006-1888  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/405,271A  
 ; FILING DATE: 14-MAR-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MORASHIGE, KATE H.  
 ; REGISTRATION NUMBER: 29,959  
 ; REFERENCE/DOCKET NUMBER: 22000-20526.22  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 887-1500  
 ; TELEFAX: (202) 887-0763  
 ; TELEX: 90-4030 MRSNFOERSH  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1981 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-405-271A-15

Alignment Scores:  
 Pred. No.: 2,31e-18 Length: 1981  
 Score: 267.50 Matches: 107  
 Percent Similarity: 42.28% Conservative: 49  
 Best Local Similarity: 29.00% Mismatches: 151  
 Query Match: 13.82% Indels: 63  
 DB: 4 Gaps: 14

US-10-023-586B-2 (1-360) x US-08-405-271A-15 (1-1981)

QY 24 GlyAlaSerArgAspMetGluLysValaspMetAsnThrSerGlnGluGlnGly----- 41  
 DB 334 GGCTCTGGCTCAACTGTGCTCCACGTTGATGGAACACAGTCCGACCCATCGGTCTCTAAC 393  
 QY 42 -----LeuCysGlnPheSer-----GluLysTyr 49  
 DB 394 CGACGGGCTTGGCGGGAACGACACGCTGTGCCCTCAGACCGGAGCCCTCCATGGTC 453  
 QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGluProLeuAsn 69  
 DB 454 ACAGCCATCACCATGCGCCCTCTATTCTGCTGTGTGAGTGGCCCTCTTTGGAAAC 513  
 QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrTyr 89  
 DB 514 TTCTCTGTCATGTATGTGTTGTAAGATATACCAAAATGAAGACTGCCACCAACATCTAC 573  
 QY 90 LeuValAsnLeuMetValAlaAspLeuTyrVal---LeuLeuProPheLeuIle 108  
 DB 574 ATTTTCAACCTGCTCTGCAGATGCTTAGCCACTAGCAGCGCTGCCCTTTCAGAGTGT 633  
 QY 109 ThrTyrSerLeuAspArgTrpProPheGlyGluLeuLysLysLeuValHisPhe 128  
 DB 634 AACTAC---CTGATGGGAAGCTGGCCCTTGGAAACATCTCTGCAAGATCGTATCTCA 690  
 QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
 DB 691 ATAGACTACTACAACATGTGTACCAGTATCTTCAACCCTCTGCACCATGAGTGTAGACCG 750  
 QY 149 PheLeuGlyValCysHisProLysSerLeuProTyrArgThrArgHisAlaTrp 168  
 DB 751 TACATTGGCTGTGCCACCGGTCAGGCGCTGGATTCTCGTACCCCGCAAGATGCCAA 810  
 QY 169 LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188  
 DB 811 ATTGTCAATGTCTGCAACTGGATCTCTCTGCGCATTTGGTCTGCCGTATGTTCATG 870  
 QY 189 SerHisThrAspTyrIleAsnGlyGlnMet-----Ile 199  
 DB 871 GCAACACACAAAATACAGCAGGCGTCCATAGATTGACCCCTCAGTCTCTCATCCACA 930  
 QY 200 TrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeu 219  
 DB 931 TGGTAC-----TGGGGAACCTGCTCAAAATCTGTCTCTTCATCTTCGCTTC 978  
 QY 220 ThrLeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAsp----- 237  
 DB 979 ATCATGCGCGGC-----CTCAT-CATCACTGTGTGTTAGTACTGATGATCTT 1025  
 QY 238 ---GlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerPro 256  
 DB 1026 ACAGCTCAAGAGTGTCCGATGCTGTGCGGCTCCAAAGAAAGAGCAGGAA-----CCT 1079  
 QY 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLysPheCysAla 276  
 DB 1080 GCGCAG-----GATCACCAGGATGTGTGGTGGTGGTGTATTATTATGCTGTG 1130  
 QY 277 LeuProTyrHisSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuPro 296  
 DB 1131 CTGGACCCCATCCATCTATGTCTATCATCAAGACACTGATCAGCATTCAGAAACCCAC 1190  
 QY 297 AlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGln 316  
 DB 1191 TTTCCAGAC-----TGTTCCTGGCACTTCTCATTTGCTTGGGTATACAAACAG 1241  
 QY 317 LeuProGlnProSerProValLeuSerPheLysGlyGlyLys----- 330  
 DB 1242 CTGCCTGAACCCAGTTCT---TTATGGTTCTCTGGATGATGAACCTCAACAGATGTTTAG 1298  
 QY 331 -----AsnArgValArgLeuLeu-GlnLysLeuArgG1 341  
 DB 1299 AGAGTTCTGCATCCCAACTTCTCCACATCGAACACGAAACTCTGCTCGATCCGTCA 1358  
 QY 341 nAsnLysLeuGlyGluHisProAla 349  
 DB 1359 AAAC---ACTAGGGAACACCCCTCC 1380

```
Oy 70 GlyThrValLeuTrpHisSerTrpGlyClnThrLysArgTrpSerCysAlaThrThrTyr 89
Db 514 TTCCTGTCATGTATGATGTAAGATATACCAAAATGAAGACTGCCACCAACATCTAC 573
Oy 90 LeuValAsnLeuMetValAlaAspLeuLeuTyrVal---LeuLeuProPheLeuIleIle 108
Db 574 ATTTTCAACCTTGCTCTGGCAGATGCTTACCCACTAGCAGCTGCCCTTTTCAGAGTGTT 633
Oy 109 ThrTyrSerLeuAspAspArgTrpPheGlyGluLeuLeuCysLysLeuValHisPhe 128
Db 634 AACTAC---CTGATGGGAACGTGGCCCTTGGAAACATCTCTCGAAGATCGTCTATCTCA 690
Oy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 691 ATAGACTACTACACATGTTCCACAGTATCTTCCCTCTGCACCATGAGTGTAGACGGC 750
Oy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrp 168
Db 751 TACATTGCCGTCCTGCCACCCCGCTCAAGCCCTGGATTCCCGTACCCCGCGAAATGCCAAA 810
Oy 169 LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188
Db 811 ATTGTCAATGCTGCAACTGGATCTCTCTCTGCCATTGGCTCTGCCCGTAATGTTCATG 870
Oy 189 SerHisThrAspTyrIleAsnGlyClnMet-----Ile 199
Db 871 GCAACCAACAAAATACAGCAGCGGGTCCATAGATTGCACCCCTCACGTTCTCTCATCCACA 930
Oy 200 TrpTyrAspMetThrSerGlnGluAspPheAspArgLeuPheAlaTyrGlyIleValLeu 219
Db 931 TGGTAC-----TGGGAGAACCTGCTCAAAATCTGTCTCTCATCTTCATCTCGCCTTC 978
Oy 220 ThrLeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAsp----- 237
Db 979 ATCATGCCGGGC-----CTCAT-CATCAGCTGTGTGTTTGGACTGATGATCTT 1025
Oy 238 ---GlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerPro 256
Db 1026 ACAGCTCAAGAGTCTCGCATGCTGCTGGGCTCCAAAGAAAAGGACAGGAA-----CCT 1079
Oy 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAla 276
Db 1080 GCGCAG-----GATCACCCGGATGGTCTGGTGGTCTGGTGGTGTATTATTGTCTG 1130
Oy 277 LeuProTyrHisSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuPro 296
Db 1131 CTGGACCCCATCCACATCTATGTCATCATCAAGCACTGATCAGATTCCAGAAACCCAC 1190
Oy 297 AlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGln 316
Db 1191 TTTCAGAC-----TGTTCTGGCACTTCTGCATTCGCTTGGGTTACACAAACAG 1241
Oy 317 LeuProGlnProSerProValLeuSerPheLysGlyLys----- 330
Db 1242 CTGCTGAACCCAGTTCT---TTATGGCTTCTCGATGAAACCTCAACCGATGTTTATG 1298
Oy 331 -----AsnArgValArgLeuLeuGlnLysLeuArgG1 341
Db 1299 AGAGTTCTGTCATCCCACTTCTCCACAATCGAAGACAAACCTCTGCTCGAATCCGTC 1358
Oy 341 nAsnLysLeuGlyGluHisProAla 349
Db 1359 AAAC---ACTAGGGAACACCCCTCC 1380
```

Search completed: May 27, 2003, 00:32:37  
Job time : 76 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:57:58 ; Search time 201 Seconds  
(without alignments)  
2365.007 Million cell updates/sec

Title: US-10-023-586b-2

Perfect score: 1936

Sequence: 1 MSLILLPSRSGSRGRRGAL.....QNKLGHPAGKRCPCLNRS 360

coring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q-/cgn2\_1/USFTO\_spool/US1023586/runat\_22052003\_114852\_28717/app\_query.fasta.1.519  
-DB-Published-Applications\_NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=ext -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -SCOR=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US1023586 -CGN.1.1.138 -runat\_22052003\_114852\_28717  
-NCPU=6 -ICPU=3 -NO\_XLPYX -NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPT=6  
-FGAPEXT=7 -YGAPOPT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1792	92.6	1017	9	US-09-885-453-3
2	921	47.6	510	9	US-09-782-974C-13
3	445.5	23.0	1429	9	US-09-077-173A-1
c 4	370	19.1	.551	9	US-09-791-932-32

Alignment Scores: 1.47e-193 Length: 1017  
Pred. No.: 1792.00 Matches: 338  
Score: 99.71%  
Percent Similarity: 99.71%  
Conservative: 0

ALIGNMENTS

RESULT 1

US-09-885-453-3  
; Sequence 3, Application US/09885453  
; Publication No. US20030088080A1  
; GENERAL INFORMATION:  
; APPLICANT: Communi, Didier  
; TITLE OF INVENTION: RECEPTOR GPCRxl0  
; FILE REFERENCE: 9409/2082  
; CURRENT APPLICATION NUMBER: US/09/885,453  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/885,453  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1017  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DNA sequence  
; LOCATION: (1)..(1017)  
; OTHER INFORMATION: GPCRxl0 DNA sequence  
US-09-885-453-3

Sequence 85, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 7, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 12, Appli  
Sequence 7, Appli  
Sequence 12, Appli  
Sequence 15, Appli  
Sequence 10, Appli  
Sequence 8, Appli  
Sequence 26, Appli  
Sequence 26, Appli  
Sequence 6, Appli  
Sequence 25, Appli  
Sequence 25, Appli  
Sequence 13, Appli  
Sequence 12, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 9, Appli  
Sequence 22, Appli  
Sequence 22, Appli  
Sequence 3, Appli  
Sequence 11, Appli  
Sequence 51, Appli  
Sequence 51, Appli  
Sequence 1, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 4, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 16, Appli  
Sequence 16, Appli  
Sequence 16, Appli  
Sequence 24, Appli

Best Local Similarity: 99.71% Mismatches: 1  
Query Match: 92.56% Indels: 1  
DB: 9 Gaps: 0

US-10-023-586B-2 (1-360) x US-09-885-453-3 (1-1017)

```
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
DB 1 ATGCTGTCCATTTTGTCTTCCAGGGAAGCAGAGGGAGCGTGTGGAGCTCTG 60
QY 21 LeuLeuGluGlyValSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
DB 61 CTCCTGGAGGAGCTTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA 120
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
DB 121 GGTCTCTGCCAGTCTTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 180
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
DB 181 ATCTTTATCTAGGCTGTCACCTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
DB 241 AAGCGCTGGAGCTGTCCACCACTATCTGGGAACCTGATGGCCGACCTGCTTTAT 300
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
DB 301 GTGCTATTGCCCTTCTCATCATCACTACTACCTAGATGACAGGTGGCCCTTCGGGGAG 360
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
DB 361 CTGCTCTGCAAGCTGTGCACTCTCTGTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 420
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
DB 421 CTGACCTGCTCTCTGTGCACCACTCTCTAGGTGTGTGCCACCACTGTGTGGCTGCC 480
QY 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeu 180
DB 481 TACCGGACCCGAGGATGCTGTGGTGGGCACGACGACCACTGGGCCCTGTGTGGTCTC 540
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
DB 541 CAGCTGTCTGCCACACTGGCCCTCTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
DB 601 TATGACATGACCCAGCAAGAGATTTTGTATCGCTTTTGTCCCTACGGCATAGTCTTGACA 660
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
DB 661 TTGTCTGGCTTTT-TCCCTCTTGTCTATTGTTGTGCTATTCTACTGATGTCAGGAG 719
QY 241 ProAspGlnAlaArgGlyCluProHisGluAspArgGlnHisSerProSerGlnValHis 260
DB 720 CTGTATCAAGCCAGAGGAGAACTCATGAGGACAGCAACAGCCGAGCCGAGGTCCTAT 779
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
DB 780 CCGGACCATCTCTACTGGGTGTGGCTCTTTCACCTCTCTTTTGTGGCTTCCATATCAC 839
QY 281 SerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
DB 840 TCGCTCTCTTCTACCTACCATCTGCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 899
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
DB 900 AGCCAGTGTGGCTTCAAGATATGAGGGCTCTGTGTAGTGTGAGCAGCTGCTCAACCC 959
QY 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
DB 960 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGATGACGCTCTCTCCAGAACTG 1016
```

RESULT 2

```
US-09-782-974C-13
; Sequence 13, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 41USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-13
```

Alignment Scores:  
Pred. No.: 3,16e-95 Length: 510  
Score: 921.00 Matches: 170  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 47.57% Indels: 0  
DB: Gaps: 0

US-10-023-586B-2 (1-360) x US-09-782-974C-13 (1-510)

```
QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102
DB 1 TGGAGCTGTGGCCACCACTATCTGTGGAACCTGTGTCGCCGACCTGCTTATGTGCTA 60
QY 103 LeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeu 122
DB 61 TTGCCCTTCTCATCATCACTACTACCTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120
QY 123 CysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThr 142
DB 121 TGCAAGCTGCTGCACTTCTGTCTATATCAACCTTTAGCGCAGCATCTGCTGTGCTGAC 180
QY 143 CysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg 162
DB 181 TGCATCTCTGTGCACAGTCTCTAGGTGTGCCACCACTGTGTGCTGCTGCTACCGG 240
QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeu 182
DB 241 ACCCGCAGGCATGCTGTGGTGGGCACCAACCACTGGGCCCTGTGTGCTCTCCAGCTG 300
```





; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,303  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,397  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,247  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/188,880  
; PRIOR FILING DATE: 2000-03-13  
; PRIOR APPLICATION NUMBER: 60/217,369  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/217,370  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/218,492  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/186,810  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/188,064  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: 60/186,457  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: 60/213,861  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/194,344  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/218,337  
; PRIOR FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 32  
; LENGTH: 551  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-791-932-32

Alignment Scores:  
Pred. No.: 1.13e-32 Length: 551  
Score: 370.00 Matches: 74  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.11% Indels: 0  
DB: 9 Gaps: 0

US-10-023-586B-2 (1-360) x US-09-791-932-32 (1-551)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
Db 222 ATGCTGTCCATTTCCTTCCTCCAGGGAGCAGAGCGGAGCGCTCGTGGAGCTCTG 163  
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
Db 162 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATCAATACATCAGAGAACAA 103  
41 GlyLeuGlnPheSerGlnLysThrLysGlnValThrLeuSerLeuAlaThrSerIle 60  
Db 102 GGTCTGCTCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGATAC 43  
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyr 74  
Db 42 ATCTTTATCTAGGCTGCCACTAATATGCACTGCTCTGTGG 1

RESULT 5

US-09-782-974C-85  
; Sequence 85, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor  
; FILE REFERENCE: 41USPHRM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C

; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/165,838  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 60/198,568  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 85  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-782-974C-85

Alignment Scores:  
Pred. No.: 3.21e-22 Length: 1020  
Score: 281.50 Matches: 79  
Percent Similarity: 49.00% Conservative: 43  
Best Local Similarity: 31.73% Mismatches: 110  
Query Match: 14.54% Indels: 20  
DB: 9 Gaps: 6

US-10-023-586B-2 (1-360) x US-09-782-974C-85 (1-1020)

Qy 50 LysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeuAsn 69  
Db 97 AAGATGCACCTACCTCCCTGTTATTTATGTCATATCTCTCTGGGATTTCCAGGCAAT 156  
Qy 70 GlyThrValLeuTyrPheSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db 157 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGAAGAGCAGACCATCAT 216  
Qy 90 LeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 108  
Db 217 ATGCTGAACCTGGCTGCACAGATCTGCTGTATCTGACCCAGCCTCCCTTCCTGATTAC 276  
Qy 109 ThrTyrSerLeuAspArgTrpPheGlyGluLeuLeuCysLysLeuValHisPhe 128  
Db 277 TACTATGCCAGTGGCGAAACTGGATCTTTGGAGATTTCATGTGTAAGTTTATCCGCTTC 336  
Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
Db 337 AGCTTCCATTTCAACCTGTATAGCAGCATCTCTCTCCACCTGTTTTCAGCATCTCCGC 396  
Qy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrp 168  
Db 397 TACTGTGTGATCATTCACCAATAGCTCTTTTCCATTCACAAAACCTCGATGTGCAATT 456  
Qy 169 LeuGlyThrSerThrThrAlaLeuValLeuGlnLeuLeuPro---ThrLeuAla 187  
Db 457 GTAGCCTGTGCTGTGGTGGATCATTTCCATGTGAGTGTGATCCGATGACCTTCTTG 516  
Qy 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleThrTyrAspMetThrSerGlnGlu 207  
Db 517 ATCAGATCAACCAACAGGACCAAC---AGATCAGCCTGTCTCGACCTCACCAGTTCGGAT 573  
Qy 208 AsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227

Db 574 GAACCTCAATACTATTAAAGTGGTACAAACCTGATTGTTGACTGCAAGTAC-TTTCGCGCTCCC 632  
QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProaspGlnAlaArgGlyGlu 247  
Db 633 CTGGGGTAGTAGTACACTTTGGCTA-----TACCACGATTAT 668  
QY 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProasp 262  
Db 669 CCACACTTTGACCCAGTGGAGCTGCAAACTGACAGCT-GCCTTAAGCAGAAAGCAGCAGGC 727  
QY 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282  
Db 728 TAACCATTC---TGCTACTCTTGGCAATT-TACGTATGTTTTTACCCCTCCCATATCTTG 783  
283 LeuLeuProHisHisLeuLeuSerAla 291  
Db 784 AGGTCATTGAGTCAATCTCAGCC 810

## RESULT 6

US-10-023-775B-1  
; Sequence 1, Application US/10023775B  
; Publication No. US2003002282A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Ltd. (EP/GB only)  
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)  
; APPLICANT: Fidock, Mark David  
; TITLE OF INVENTION: No. US2003002282A1el Polypeptide  
; FILE REFERENCE: PC10959AGPR  
; CURRENT APPLICATION NUMBER: US/10/023,775B  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: GB 0030854.4  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/260,590  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: US 60/296,660  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: GB 0111031.1  
; PRIOR FILING DATE: 2001-05-04  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1014  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-023-775B-1

Alignment Scores:  
Pred. No.: 6,95e-22 Length: 1014  
Score: 278.50 Matches: 76  
Percent Similarity: 49.17% Conservative: 42  
Best Local Similarity: 31.67% Mismatches: 105  
Query Match: 14.39% Indels: 20  
DB: 9 Gaps: 6

US-10-023-586B-2 (1-360) x US-10-023-775B-1 (1-1014)

QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIlePheLeuGlyLeuProLeuAsn 69  
Db 94 AAGATGCACTACCTCCCTGTTATTATGGCATTTATCTCTCGTGGGATTTCCAGGCAAT 153  
QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGGAGAGCAGCACCACCATCATT 213  
QY 90 LeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 108  
Db 214 ATGCTGAACCTGGCCTGCACAGATCTGCTGTATCTGACACCGCTCCCTCCCTGATTTCAC 273  
QY 109 ThrTyrSerLeuAspArgTrpProPheGlyGlnLeuLeuCysLysLeuValHisPhe 128  
Db 274 TACTATCCCAAGTGGCGAAACTGGATCTTTGGAGATTTCATGTGAAGTTTATCCCGCTTC 333

QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
Db 334 AGCTTCCATTCAACCTGTATAGCAGCATCTCTCTCTACCTGTTTCACGATCTTCCGCG 393  
QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTrp 168  
Db 394 TACTGTGTGATCATTCACCAATGAGCTGCTTTTCCATTCCAAAACTCGATGTGCAGTT 453  
QY 169 LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuPro---ThrLeuAla 187  
Db 454 GTAGCTGTGCTGTGGTGTGGATCAITTCACCTGTCTCATTCCTCGATGACCTTCTTG 513  
QY 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGlu 207  
Db 514 ATCACAATCAACCAACAGGACCAAC---AGATCAGCCTGTCTCGACCTCACCAGTTCGAT 570  
QY 208 AsnPheAspArgLeuPheAlaTyrGlyValLeuThrLeuSerGlyPheLeuSerLeu 227  
Db 571 GAACCTCAATACTATTAAAGTGGTACAACTAATTTTGACTGCAACTAC-TTTCGCTCC 629  
QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProaspGlnAlaArgGlyGlu 247  
Db 630 CTGGGTGATAGTGACACTTTGCTA-----TACCACGATTAT 665  
QY 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProasp 262  
Db 666 CCACACTCTGACCCAGTGGAGCTGCAAACTGACAGCT-GCCTTAAGCAGAAAGCAGGAGGC 724  
QY 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282  
Db 725 TAACCATTC---TGCTACTCTCTGCAATT-TACGTATGTTTTTACCCCTCCCATATCTTG 780

## RESULT 7

US-10-270-144-1  
; Sequence 1, Application US/10270144  
; Publication No. US200300049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000750CON  
; CURRENT APPLICATION NUMBER: US/10/270,144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1014  
; TYPE: DNA  
; ORGANISM: Human  
US-10-270-144-1

## Alignment Scores:

Pred. No.: 6,95e-22 Length: 1014  
Score: 278.50 Matches: 76  
Percent Similarity: 49.17% Conservative: 42  
Best Local Similarity: 31.67% Mismatches: 105  
Query Match: 14.39% Indels: 20  
DB: 9 Gaps: 6

US-10-023-586B-2 (1-360) x US-10-270-144-1 (1-1014)

QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIlePheLeuGlyLeuProLeuAsn 69  
Db 94 AAGATGCACTACCTCCCTGTTATTATGGCATTTATCTCTCGTGGGATTTCCAGGCAAT 153  
QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGGAGAGCAGCACCACCATCATT 213  
QY 90 LeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 108

```

214 ATGCTGAACCTGGCGCTGCACAGATCTGTATCTGCACGACCTCCCTCTCTGATTAC 273
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 109 ThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe 128
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 274 TACTATGCGACGTGGCGAAACCTGGATCTTTGGAGATTTTCATGTGTAAGTTTATCCGCTTC 333
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuLeuCysIleSerValHisGln 148
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 334 AGCTTCCATTTCAACCTGTATAGCAGATCTCTCTCCCTCAGCTGTTTCCAGCATCTTCGCG 393
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgTrpArgArgHisAlaTrp 168
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 394 TACTGTGTGATCATTCACCAATGAGCTGCTTTTCCATTCACAAACTCGATGTGCAGTT 453
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 169 LeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuPro--ThrLeuAla 187
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 454 GTAGCTGTGTGTGGTGGATCATTTACATGGTAGCTGATTCCTCCGATGACCTTCTTG 513
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGlu 207
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 514 ATCATCATCAACCAACAGACCAAC---AGATCAGCTGTCTCGACCTCACCAGTTTCGGAT 570
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 208 AsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 571 GAACCTCAATACTATTAAAGTGTGTACAACTGATTTTGACTGCACACTAC-TTCTGCTCCTCC 629
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 630 CTGTGTGTATGACACTTTTGCTA-----TACCACGATTAT 665
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProAsp 262
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 666 CCACACTCTGACCCATGAGCTGCAAACTGACAGCT-GCCTTAAGCAGAAAGCAGCAAGGC 724
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 725 TAACCATTC---TGCTACTCTCTGCATTT-TACGATGTTTTTTTACCCTTCATATCTTG 780
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::

```

RESULT 8  
US-10-188-405-7

```

; Sequence 7, Application US/10188405
; Publication No. US20030082585A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Dai, Kang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Zhao, Jiagang
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: NO. US20030082585A1el Receptors
; FILE REFERENCE: 018781-008410US
; CURRENT APPLICATION NUMBER: US/10/188,405
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/302,800
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TGR164
US-10-188-405-7

```

```

Alignment Scores:
Pred. No.: 6.95e-22 Length: 1014
Score: 278.50 Matches: 76
Percent Similarity: 49.17% Conservative: 42
Best Local Similarity: 31.67% Mismatches: 105
Query Match: 14.39% Indels: 20
DB: 9 Gaps: 6

```

```

US-10-023-586B-2 (1-360) x US-10-188-405-7 (1-1014)
Qy 50 LysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheLeuLeuGlyLeuProLeuAsn 69
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 94 AAGATGCACTACCTCCCTGGTATTATTTATGCAATATCTTCCCTCGTGGATTTTCAGGCAAT 153
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 70 GlyThrValLeuTyrHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGACACCTTGGAGAGCAGCACCATT 213
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 90 LeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIlele 108
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 214 ATGCTGAACCTGGCGCTGCACAGATCTGTGTATCTGACACGCTCCCTCTCTGATTAC 273
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 109 ThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe 128
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 274 TACTATGCGACGTGGCGAAACCTGGATCTTTGGAGATTTTCATGTGTAAGTTTATCCGCTTC 333
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuLeuCysIleSerValHisGln 148
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 334 AGCTTCCATTTCAACCTGTATAGCAGATCTCTCTCCCTCAGCTGTTTCCAGCATCTTCGCG 393
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgTrpArgArgHisAlaTrp 168
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 394 TACTGTGTGATCATTCACCAATGAGCTGCTTTTCCATTCACAAACTCGATGTGCAGTT 453
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 169 LeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuPro--ThrLeuAla 187
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 454 GTAGCTGTGTGTGGTGGATCATTTACATGGTAGCTGATTCCTCCGATGACCTTCTTG 513
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGlu 207
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 514 ATCATCATCAACCAACAGACCAAC---AGATCAGCTGTCTCGACCTCACCAGTTTCGGAT 570
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 208 AsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 571 GAACCTCAATACTATTAAAGTGTGTACAACTGATTTTGACTGCACACTAC-TTCTGCTCCTCC 629
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 630 CTGTGTGTATGACACTTTTGCTA-----TACCACGATTAT 665
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProAsp 262
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 666 CCACACTCTGACCCATGAGCTGCAAACTGACAGCT-GCCTTAAGCAGAAAGCAGCAAGGC 724
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Qy 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 725 TAACCATTC---TGCTACTCTCTGCATTT-TACGATGTTTTTTTACCCTTCATATCTTG 780
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::

```

## RESULT 9

```

US-09-885-453-2
; Sequence 2, Application US/09885453
; Publication No. US20030088080A1
; GENERAL INFORMATION:
; APPLICANT: Communi, Didier
; TITLE OF INVENTION: RECEPTOR GPCRxl0
; FILE REFERENCE: 9409/2082
; CURRENT APPLICATION NUMBER: US/09/885,453
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/885,453
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA nucleotide sequence
; LOCATION: (1)-(1014)
; OTHER INFORMATION: GPCRxl0 DNA sequence

```





Qy 208 AsnProHisGluasp 251  
Db 1117 GAACCTCAATACTAATTAAGTGTACAACTAATTTGACTGCACTACTTTGCTCCCTCCCTT 1176  
Qy 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247  
Db 1177 GGTGATAGTACACTTTGCTATACCAGAT-----TAT 1209  
Qy 248 ProHisGluasp 251  
Db 1210 CCACACTCTGAC 1221

RESULT 13

US-10-024-494-7

Sequence 7, Application US/10024494

Publication No. US20030044898A1

GENERAL INFORMATION:

APPLICANT: LI, YI

CAO, LIANG

ROSEN, CRAIG A.

TITLE OF INVENTION: Human GPR4 G-Protein Coupled Receptor and

Nucleotides Encoding Same (As Amended)

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., NW, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

OPERATING SYSTEM: IBM PC COMPATIBLE

SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/024,494

FILING DATE: 21-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/465,973

FILING DATE: 06-JUN-1995

APPLICATION NUMBER: PCT/US95/04079

FILING DATE: 30-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: STERNE, ERIC K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1140001/EKS/HCC

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1301 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 161..1192

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-024-494-7

Alignment Scores:

Pred. No.: 2.55e-21

Score: 275.00

Percent Similarity: 47.72%

Best Local Similarity: 30.88%

Query Match: 14.20%

DB: 9

Gaps: 10

Length: 1301

Matches: 88

Conservative: 48

Mismatches: 106

Indels: 44

Gaps: 10

RESULT 14

US-09-761-962-12

Sequence 12, Application US/09761962

Patent No. US20020077285A1

GENERAL INFORMATION:

APPLICANT: Memorial Sloan-Kettering Cancer Center

TITLE OF INVENTION: Identification and Characterization of Multiple Splice

Variants of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-

Variant of Mu-





Best Local Similarity: 29.00% Mismatches: 151  
Query Match: 13.82% Indels: 63  
DB: 10 Gaps: 14

US-10-023-586B-2 (1-360) x US-09-823-114-15 (1-1981)

```
Qy 24 GlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGlnGly----- 41
Db 334 GGCTCTGGCTCAACTGTCCACGTTGATGGAACAGTCCGACCATCGGTCCTAAC 393
Qy 42 -----LeuCysGlnPheSer-----GluLysTyr 49
Db 394 CCGACGGCCCTGGCGGAACGACAGCTGTGCCCTCAGACGCGCACCCCTCCATGGTC 453
Qy 50 LysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheLeuGlyLeuProLeuAsn 69
Db 454 ACAGCCATCACCATCAGCCCTCTATTCATGTGTGTGTAGTGGGCCCTCTTTGGAAAC 513
Qy 70 GlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89
Db 514 TTCCTGTCATGTATGTGATTGTAAGATATACCAAAATGAAGACTGCCACCAACATCTAC 573
Qy 90 LeuValAsnLeuMetValAlaAspLeuLysTyrVal---LeuLeuProPheLeuIleIle 108
Db 574 ATTTTCAACCTTGTCTGGCAGATGCTTACCCACTAGCAGCTGCCCTTTTCAGAGTGT 633
Qy 109 ThrTyrSerLeuAspAspArgTrpProPheGlyLeuLeuLysLysLeuValHisPhe 128
Db 634 AACTAC---CTGATGGGAACGTGGCCCTTTGGAACATCTCTGCCAAGATCGTGATCTCA 690
Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 691 ATAGACTACTACAACTGTTCCACGATGCTTCCACCTCTGCACCATGAGTGTAGACCGC 750
Qy 149 PheLeuGlyValCysHisProLeuLysSerLeuProTyrArgThrArgArgHisAlaTrp 168
Db 751 TACATTGCCGTGCCACCGCTCCAGCCCTGGATTTCGGTACCCCGGAAATGCCAAA 810
Qy 169 LeuGlyThrSerThrTrpAlaLeuValLeuGlnLeuLeuProThrLeuAlaPhe 188
Db 811 ATTGTCAATGTCGCACTGATCTCTCTCTGCCATTGGTCTGCCGTAATGTTTCATG 870
Qy 189 SerHisThrAspTyrIleAsnGlyGlnMet-----Ile 199
Db 871 GCAACCAACAAATACAGCGAGGGTCCATAGATTGCACCTCTCATCTCTCATCCACA 930
Qy 200 TrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyLeuVal 219
Db 931 TGGTAC-----TGGGAGAACCCTGCTCAAAATCTGTCTTCATCTCGCCTTC 978
Qy 220 ThrLeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAsp----- 237
Db 979 ATCATGCCGGC-----CTCAT-CATCAGCTGTGTGTATGGACTGTGATGCTT 1025
Qy 238 ---GlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerPro 256
Db 1026 ACAGCTCAAGAGTGTCCGCATGCTCTCGGGCTCCAAAGAAAAGGACAGGAA-----CCT 1079
Qy 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAla 276
Db 1080 GCGCAG-----GATCACCCGGATGGTGTGGTGTGGTGTGGTGTATTTATTTGTCG 1130
Qy 277 LeuProTyrHisSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuPro 296
Db 1131 CTGGACCCCATCCACATCTATGTCATCATCAAGCACTGATCAGCTCCAGAACCCAC 1190
Qy 297 AlaLeuAspGlySerGlnCysGlyGlnAspMetGluAlaSerGlyGluCysGluGln 316
Db 1191 TTTCCAGAC-----TGTTTCTGGCACTTCTGCACTTCTGCTGGTGTACACAAACAG 1241
Qy 317 LeuProGlnProSerProValLeuSerPheLysGlyLys----- 330
Db 1242 CTGCTGAACCCAGTTCT---TTATGCGTCTCTGGATGAAACTTCAACGATGTTTATG 1298
```

```
Qy 331 -----AspArgValArgLeuLeu-GlnLysLeuArgG1 341
Db 1299 AGAGTTCTGCATCCCAACTTCTCTCCACAATCGACAGCAAAACTCTGCTCGAATCCGTCA 1358
Qy 341 nAsnLysLeuGlyGluHisProAla 349
Db 1359 AAAC---ACTAGGAGACACCCCTCC 1380
```

Search completed: May 27, 2003, 00:36:10  
Job time : 216 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:53:18 ; Search time 1772 Seconds  
(without alignments)  
3290.279 Million cell updates/sec

Title: US-10-023-586B-2  
Perfect score: 1936  
Sequence: 1 MSLILLPSRSGSRGAL.....QNKLGHPAGKRCPLNRS 360

coring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_B2n.model -DEV=xlh  
-Q/cgn2\_1/USPto\_Spool/US10023586/runat\_22052003\_114851\_28657/app\_query.fasta\_i.519  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023586 -ACGN\_1\_1\_1906 @runat\_22052003\_114851\_28657 -NCPU=6 -ICPU=3  
-NO\_XLPYX -NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estlin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	414.5	21.4	2542	11	AK017378	AK017378 Mus muscu
2	414.5	21.4	3001	11	AK005013	AK005013 Mus muscu
3	378	19.5	851	13	BI833118	BI833118 603090834
4	360.5	18.6	641	14	BQ396255	BQ396255 NISC_pg19
5	359.5	18.6	638	9	AL675845	AL675845 AL675845
6	357	18.4	606	17	AZ953874	AZ953874 2M0219L17
7	355.5	18.4	1063	14	BM918491	BM918491 AGENCOURT
8	341	17.6	884	9	AL525099	AL525099 AL525099
9	337.5	17.4	898	13	BI819396	BI819396 603034571
c 10	325	16.8	1101	17	CNS05325	AL318925 Tetraodon
11	318	16.4	934	13	BI768397	BI768397 603053844
12	317	16.4	623	14	BQ038875	BQ038875 pgn1c.pk0
13	314	16.2	936	9	AL520218	AL520218 AL520218
14	310	16.0	801	9	AL549441	AL549441 AL549441
15	309.5	16.0	507	12	BF938665	BF938665 fm79g12.y
16	302	15.6	1118	14	BM918898	BM918898 AGENCOURT
17	300.5	15.5	899	14	BQ959110	BQ959110 AGENCOURT
18	295.5	15.3	484	14	BQ551383	BQ551383 H4008H11-
19	290	15.0	955	9	AL521440	AL521440 AL521440
c 20	286.5	14.8	408	13	BI401676	BI401676 MI-P-CP0-
21	276.5	14.3	589	14	BQ396804	BQ396804 NISC_pg23
22	274.5	14.2	781	12	BG681003	BG681003 602628822
23	272	14.0	546	12	BG552112	BG552112 ddi7a03.y
24	271.5	14.0	784	13	BI820686	BI820686 603034472
25	270	13.9	766	9	AJ450835	AJ450835 AJ450835
26	269	13.9	808	9	AJ456719	AJ456719 AJ456719
27	267	13.8	749	9	AJ447002	AJ447002 AJ447002
28	266.5	13.8	1083	14	BM805382	BM805382 AGENCOURT
29	266	13.7	747	9	AJ446611	AJ446611 AJ446611
30	266	13.7	773	9	AJ451396	AJ451396 AJ451396
31	266	13.7	798	9	AJ451155	AJ451155 AJ451155
32	266	13.7	900	14	BQ724849	BQ724849 AGENCOURT
33	265.5	13.7	806	17	CNS03VJW	AL262517 Tetraodon
34	265	13.7	800	9	AJ452624	AJ452624 AJ452624
35	265	13.7	860	9	AJ446632	AJ446632 AJ446632
36	265	13.7	890	9	AJ456561	AJ456561 AJ456561
37	264	13.6	741	9	AJ453402	AJ453402 AJ453402
38	263	13.6	740	9	AJ450921	AJ450921 AJ450921
39	262.5	13.6	845	9	AJ456135	AJ456135 AJ456135
40	262.5	13.6	879	14	BQ933137	BQ933137 AGENCOURT
41	260	13.4	1136	14	BM906977	BM906977 AGENCOURT
42	259	13.4	724	9	AJ447835	AJ447835 AJ447835
43	259	13.4	1064	13	BI517798	BI517798 603042051
44	258.5	13.4	642	12	BG159797	BG159797 de88h09.y
45	258	13.3	746	9	AJ450259	AJ450259 AJ450259

ALIGNMENTS

RESULT:1  
AK017378

LOCUS  
DEFINITION

AK017378 2542 bp mRNA linear HTC 19-JAN-2002  
Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15:purinergic receptor P2Y, G-protein coupled 2, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

AK017378.1

Gr:12856588

HTC; CAP trapper.

Mus musculus (strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library

clone:5430432J15.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

## REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

## REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## REFERENCE

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Momberts, P., Nordone, P.,  
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyooka, K., Wang, K. H., Westz, C., Whittaker, C., Wilming, L.,  
Yoshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
and Hayashizaki, Y.

## REFERENCE

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

## REFERENCE

5 (bases 1 to 2542)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.

## REFERENCE

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

## COMMENT

Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 10.0 and subtraction to  
Rot = 100.0. Second strand cDNA was prepared with the primer  
adaptor of sequence [5',  
GAGAGAGAGATCTCGAGTTAATTAATCCGCCCCCCC 3']. cDNA was cleaved  
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after  
bulk excision from lambda FLC I. Cloning sites, 5' end: Sall; 3',  
end: BamHI. Host: DH10B.

## FEATURES

Location/Qualifiers  
1..2542  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:5430432J15"  
/db\_xref="MGI:1897121"  
/db\_xref="taxon:10090"  
/clone="5430432J15"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="6 days neonate"  
1..2542  
/gene="P2ry2"  
484..1605  
/gene="P2ry2"  
/note="data source: MGD, source key: MGI:105107,  
evidence: ISS  
purinergic receptor P2Y, G-protein coupled 2  
putative"  
/codon\_start=1  
/protein\_id="BAB30719.1"  
/db\_xref="GI:12856589"  
/db\_xref="MGI:105107"  
/translation="MAADLEPNSTINGTWEGDELGYKCRNEDEKYLPLVSYGVVC  
VLGCLNVALYIFLCRLKTNASTTYMFLAVSDLSLAASPLPLVLYYARGDHPFVS  
TLCKLVLFYTNLYCSILFTCSIVHRCGLVRLPSLRGRARYARRAAVWVVL  
VLCAOPVLYFTSVTRTCHDTSARELFSEHFAVSSVLMGLFLFAVPESVILVCY  
VLMARLLPAYGTGGLPRAKRSVRTIALVLAVALCFPLPHLTRLYSFRSLDL  
SCHTLNMAVKITRPLASNSCLDPVLYFLAGORLYRFRADARPPTEPTSPQARR  
KLGHRPNRTVRKDLVSVDSSRRTESTPAGSETKDIRL"  
BASE COUNT 517 a 737 c 705 g 582 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1..le-33 Length: 2542  
Score: 414.50 Matches: 132  
Percent Similarity: 48.66% Conservative: 50  
Best Local Similarity: 35.29% Mismatches: 155  
Query Match: 21.41% Indels: 38  
DB: 11 Gaps: 9

US-10-023-586B-2 (1-360) x AK017378 (1-2542)

QY 2 LeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeuLeu 21  
|||||||  
Db 435 CTGAGCATCCTC-----ACCACTCAAGACGAGCAGGTGATCAG 473  
QY 22 LeuGluGly-AlaSerArgAspMetGluLysValaspMetAsnThrSer----- 37  
|||||||  
Db 474 GTCCAGGGCAATGGCAGCAGACCTGGAACTCCATCACCACCTGGGGA 533  
QY 38 ----GlnGluGlnGlyLeu---CysGlnPheSerGluLysTyrLysGlnValTyrLeuSe 55  
|||||||  
Db 534 GGGGCGACGAACCTGGGATACAGTGTCTTCAACGAGGACTTCAAGTACTGCTGTGCC 593





```

RESULT 3
BI833118      851 bp      mRNA      linear      EST 04-OCT-2001
LOCUS        603090834F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5229633 5',
DEFINITION   mRNA sequence.
ACCESSION    BI833118
VERSION      BI833118.1 GI:15944668
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 851)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11578 row: c column: 10
            High quality sequence stop: 847.
            Location/Qualifiers
            1. .851
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:5229633"
               /clone_lib="NIH_MGC_120"
               /lab_host="DH10B"
               /note="Organ: pooled pancreas and spleen; Vector:
               PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
               source anonymous pool of spleen and pancreas from 28 yo
               male. Library is oligo-dT primed and directionally cloned
               (EcoRV site is destroyed upon cloning). Average insert
               size 1.5 kb, insert size range 1-2.5 kb. Library is
               normalized and enriched for full-length clones and was
               constructed by C. Gruber (in vitro). Research Genetics
               tracking code 025. Note: This is a NIH_MGC Library."
FEATURES             source
BASE COUNT          127 a      311 c      231 g      182 t
ORIGIN
Alignment Scores:
Pred. No.:          1.97e-30      Length:      851
Score:              378.00      Matches:    83
Percent Similarity: 52.76%      Conservative: 22
Best Local Similarity: 41.71%      Mismatches: 88
Query Match:        19.52%      Indels:     6
DB:                  13          Gaps:       3

US-10-023-586B-2 (1-360) x BI833118 (1-851)
Qy 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluIlys 48
Db 116 GACATGGCACAGCGCAGGCTCTGGCTTGCACCCACCACTGTGTCTACCGCGAGAAC 175
Qy 49 TyrLysGlnValTyrSerGlnAlaTyrSerIlellePheLeuGlyLeuProLeu 68
Db 176 TTCAAGCAACTGCTGCTGCCACCTGTGTATTCGGCGGTGGCGGCTGCGCTCCGCTG 235
Qy 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThr 88
Db 236 AACATCTGTGTCAATCCAGATCTGCACGTCGCGCGGCGCTGCGCCGACGCGCGT 295
Qy 89 TyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeu 107
Db 296 TACACCTAAACCTTGCTCTGGCTGACCTGCTATATGCTGCTCCCTGCCCTGCTCATC 355

```

---

```

Qy 108 IleThrTyrSerLeuAspArgTirProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 356 TACAACTATGCCCAAGTGATCACTGGCCCTTTTGGGACTTCGGCTGGCGCTGGTCCGC 415
Qy 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 416 TTCCTCTTCTATGCAACACCTGCAGCGAGCATCTCTTCTCTACCTGCATCAGCTTCAG 475
Qy 148 GlnPheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrArg---ThrArgArgHis 166
Db 476 CGCTACCTGGGCATCTGCCACCCCTGGCCCTGGCCCAACACCTGGGGGCGCCGGCT 535
Qy 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186
Db 536 GCCTGGCTAGTGTGTAGCGGTGTGGCTGGCGGTGACACCCAGTGCCTGCCACAGCC 595
Qy 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206
Db 596 ATCTTCGCTGCCACAGGCATCCAGCAACCGCACTGTCTGTATGACCTCAGCCGCGCT 655
Qy 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu 225
Db 656 GCCCTGCCACCCACTATATGCCCTATGCGCATGGCTCTCACTGTCACTGCGCTTCTCG 712
RESULT 4
BO396255      641 bp      mRNA      linear      EST 22-MAY-2002
LOCUS        NISC-ng19g07.y1 NICHG XGC Emb6 Silurana tropicalis cDNA clone
DEFINITION   IMAGE:5383884 5', mRNA sequence.
ACCESSION    BO396255
VERSION      BO396255.1 GI:21083932
KEYWORDS     EST.
SOURCE       western clawed frog.
ORGANISM     Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE    1 (bases 1 to 641)
AUTHORS     NIH-XGC http://image.llnl.gov/image/html/xenopuslib.info.shtml.
TITLE       National Institute of Child Health and Human Development, National
            Cancer Institute, Xenopus Gene Collection
JOURNAL     Unpublished (2002)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            cDNA Library Preparation:
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM11977 row: N column: 13
            Seq primer: M13Rpl reverse primer (ABI).
FEATURES             source
BASE COUNT          132 a      184 c      153 g      172 t
ORIGIN
Alignment Scores:
Pred. No.:          1.03e-28      Length:      641
Score:              360.50      Matches:    77
Percent Similarity: 55.61%      Conservative: 32

```









```

Db 569 TACACCCCTAAACCTTGCTCTGGCTGACCTGATATGCTGCTCCCTCCCTGCTCATC 628
Oy 108 lIleThrTyrSerLeuAspArgTrrpPheGlyGluLeuLeuCysLysLeuValHis 127
Db 629 TACAACTATGCCCAAGGTGATCACTGGCCCTTTGGGAGCTTCGCTGCGGCTGGTCCGC 688
Oy 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 689 TTCCTCTTCTATGCAACCTGCACGGCAGCATCTCTCTCCCTCACTGCATCAGCTTCCAG 748
Oy 148 GlnPheLeuValCysHisProLeuLeuCysSerLeuProTyrArg---ThrArgHis 166
Db 749 CGTACCTGGGCACTGCACCGCTGGCCCTGGCCGACCAACCTGGGGCGCGGGCT 808
Oy 167 AlaTrpLeuGlyThrSerThrThrAlaLeuValLeuGlnLeuLeuProThrLeu 186
Db 809 GCCTGGCTAGTGTGTAGCGGTGGCTGGCGGTGACCAACCTGCTGCGCCACAGCC 868

RESULT 9
BI819396 898 bp mRNA linear EST 04-OCT-2001
LOCUS 603034571f1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175761 5',
DEFINITION mRNA sequence.
ACCESSION BI819396
VERSION BI819396.1 GI:15930946
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 898)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11437 row: n column: 18
High quality sequence stop: 884.
Location/Qualifiers
1. 898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175761"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
149 a 303 c 255 g 191 t
BASE COUNT
ORIGIN

Alignment Scores: 5.51e-26 Length: 898
Pred. No.: 337.50 Matches: 83
Score: 51.96%
Percent Similarity: Conservative: 23

```

```

Best Local Similarity: 40.69% Mismatches: 87
Query Match: 17.43% Indels: 12
DB: 13 Gaps: 4
US-10-023-586B-2 (1-360) x BI819396 (1-898)
Oy 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48
Db 201 GACAATGGCACAGGCGAGCTCTGGGCTTGGCACCCACCACTGCTGTCTACCGCGAGAAC 260
Oy 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68
Db 261 TTCAGCAACCTGCTGCTGCCACCTGTGTATTGGGGGTGCTGGCGGCTGGCTGCGCGT 320
Oy 69 AsnGlyThrValLeuLeuTrpHisSerTrpGlyGlnThrLysAlaGTrpSerCysAlaThrThr 88
Db 321 AACATCTGTGTATACCCAGATCTGCAGCTCCCGCGGGCGCTGACCGGACGCGCGCTG 380
Oy 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
Db 381 TACACCCCTAAACCTTGCTCTGGCTGACCTGTATATGCTGCTCCCTGCCCTGCTCATC 440
Oy 108 lIleThrTyrSerLeuAspArgTrrpPheGlyGluLeuLeuCysLysLeuValHis 127
Db 441 TACAACCTATGCCCAAGGTGATCACTGGCCCTTTGGGAGCTTCGCTGCGGCTGGTCCGC 500
Oy 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 501 TTCCTCTTCTATGCAACCTGCACGGCAGCATCTCTCTCCCTCACTGCATCAGCTTCCAG 560
Oy 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArg----- 164
Db 561 CGCTACCTGGGCACTGCGCACCCGCTGCC-----CCCTGGCACAAACGTGGGGGCGCT 614
Oy 165 -ArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValLeuGlnLeuLeuPr 184
Db 615 CGGGCTCGCTGGTGTGTGTAGCGGTGGTGGTGGCGGTGACAAACCACTGCTGCC 674
Oy 184 oThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrrpTyrAspMeth 204
Db 675 CACAGCCATCTTGGTGGCCACAGGATCCAGGTAACCGCACTGCTGCTGATGACTCA 734
Oy 204 rSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuLeuSerGlyph 224
Db 735 CGCGCTCGCTG-GGACCCCACTATATGCCCTATGGCATGGCTCTCTCATGTCGCTT 793
Oy 224 eLeuSerLeu 227
Db 794 CTGGTGCCTT 803

RESULT 10
CNS05325/c 1101 bp DNA linear GSS 26-JUL-2000
LOCUS . Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 020M21 of library A from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL318925
VERSION AL318925.1 GI:9551809
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 1101)
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633

TITLE
JOURNAL
MEDLINE

```



Score: 317.00 Matches: 69  
Percent Similarity: 58.60% Conservative: 23  
Best Local Similarity: 43.95% Mismatches: 63  
Query Match: 16.37% Indels: 2  
DB: 14 Gaps: 2

US-10-023-586b-2 (1-360) x B0038875 (1-623)

QY 53 TLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrVal 72  
Db 135 TACTGCCACCGTCTACATCTCGTCTTATCACCAGGCTTCTGGGCACACAGCGTGGCC 194  
QY 73 LeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsn 92  
Db 195 ATCTGGATGTCGTCTTCCACATCGCGCGTGGAGCGCATCTCGGTGTACATGTTTCAAC 254  
QY 93 LeuMetValAlaLeuLeuTyrValLeu---LeuProPheLeuIleLeuThrTyrSer 111  
Db 255 CTGGCTCTGGCCGACTTCCCTGTATGCTCCTCAGCGTCCGCCCTCATCTTCTACTACTTC 314  
QY 112 LeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyr 131  
Db 315 AACAAACCCACTGGATCTTCGGGGAGCTCATGTGCAAGTGCAGAGTTTCATTTTCCAC 374  
QY 132 IleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGly 151  
Db 375 GTGAACCTCTACGGCAGCATCTTCTCCTCAGCTGCATAAGCGTGCACAGGTACACGGGC 434  
QY 152 ValCysHisProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrpLeuGlyThr 171  
Db 435 GTCTGTGACCGCTGAAGTCTCGTGGGAGGCTGAAGAAGAAGACGCGTGTACGTGACG 494  
QY 172 SerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThr 191  
Db 495 TCCTGTGTGGCCCTGTGGTGGCGCTCATCGCCGCCATCTCTTCTACTCAGGACG 554  
QY 192 AspTyrIleAsnGlyGlnMetIle---TrpTyrAspMetThrSerGlnGlu 207  
Db 555 NNNGTAGGAGGAACAAACCATCACCTGCTACGACACACCGCGNGACGAG 605

RESULT 13  
AL520218 936 bp mRNA linear EST 13-FEB-2001  
LOCUS AL520218 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB006VE03 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL520218 GI:12783711  
VERSION AL520218.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 936)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source  
1. .936  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DB006VE03"  
/clone\_lib="LTI\_NFL004\_NBC2"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6

QY 108 IleThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127  
Db 511 TACAACATATGCCCAAGGTGATCACTGGCCTTTGGCGACTTCGCTGCCCGCTGGTCCGC 570  
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147  
Db 571 TTCTCTTCTATGCACACCTGCACCGGAGCATCTCTTCTCCTACCTGCATCAGCTTCCAG 630  
QY 148 GlnPheLeuGlyValCysHisProLeuCysSer-LeuProTyrArgThrArgArgHisAl 167  
Db 631 CGTACTGGCGCATCTCCACCGCTGGCCCTGGGCACAAACGTGGGGCGCGGGCTGC 690  
QY 167 atrpLeuLeuThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAl 187  
Db 691 CTGGCTAGTGTGTAGCCCGTGGTGGCCCGTGACAAACAGTGCCTGCCACAGCCAT 750  
QY 187 aPheSerHisThrAspTyrIleAsn---GlyGlnMetIleTrpTyrAspMetThrSerGl 206  
Db 751 CTTTGGCTGCACAGG-CATTCAAGCGTAAACGAGCTGCTGTGTATGACCTGACCTGCGTC 809  
QY 206 nGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeu-----ThrLeuSe 222  
Db 810 TGG-CCTGGGCAACCACTATGCTATGCGATGGCGCTCCACTGGTGCATCGGTCTTGGC 868  
QY 222 rGlyPheLeuSerLeuLeuGlyHisPheGly 232  
Db 869 TGCCTTTGCTGCCCTGTGGCTGCTACTGCG 899

RESULT 12  
B0038875 623 bp mRNA linear EST 01-MAY-2002  
LOCUS B0038875  
DEFINITION pgnlc.pk010.113 normalized chicken lymphoid cDNA library Gallus  
gallus cDNA clone pgnlc.pk010.113 5', similar to  
SP|P34996|P2YR.CHICK.P2Y.PURINOCEPTOR.1 (ATP RECEPTOR) (P2Y1)  
(PURINERGIC RECEPTOR) pir|S33733 G protein-coupled receptor  
chicken emb|CAA51716.1| (X73268) ATP receptor P2Y1 [Gallus gallus],  
mRNA sequence.  
ACCESSION B0038875  
VERSION B0038875.2 GI:20383637  
KEYWORDS EST.  
SOURCE chicken.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 623)  
AUTHORS Morgan, R.W. and Burnside, J.  
TITLE Chicken lymphoid ESTs  
JOURNAL Unpublished (2001)  
COMMENT On Mar 27, 2002 this sequence version replaced gi:19772415.  
Contact: Joan Burnside  
Molecular Endocrinology  
University of Delaware  
40 Townsend Hall, Newark, DE 19717, USA  
Tel.: 302 831-1345  
Fax: 302-831-3411  
Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES  
source  
1. .623  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="pgnlc.pk010.113"  
/clone\_lib="normalized chicken lymphoid cDNA library"  
/sex="Male and Female"  
/tissue\_type="thymus, bursa, spleen, PBL, bone marrow"  
/lab\_host="E.coli EMDH10B"  
/note="Vector: pCMVSPORT 6"  
BASE COUNT 117 a 199 c 172 g 128 t  
ORIGIN

Alignment Scores: 5.41e-24 Length: 623  
Pred. No.:





## BF938665

LOCUS BF938665 507 bp mRNA linear EST 22-JAN-2001  
 DEFINITION fm79q12.v1 zebrafish gridded kidney Dario rerio cDNA clone 4234174  
 5' similar to TR:057466 057466 G PROTEIN COUPLED P2Y NUCLEOTIDE  
 RECEPTOR.; mRNA sequence.

ACCESSION BF938665  
 VERSION BF938665.1 GI:1235985  
 KEYWORDS EST.  
 SOURCE zebrafish.

## ORGANISM

Dario rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
 ; Cyprinidae; Dario.

## REFERENCE

1 (bases 1 to 507)  
 Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy  
 S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood  
 K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,  
 Waller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,  
 Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.  
 and Wilson, R.

WashU Zebrafish EST Project 1998

## TITLE

Unpublished (1998)

## JOURNAL

Contact: Stephen L. Johnson

Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@watson.wustl.edu

CNA Library Preparation: Leonard Ira Zon DNA Sequencing by:  
 Washington University Genome Sequencing Center Clone distribution:  
 Genome Systems, St. Louis, Missouri (web address:  
 www.genomesystems.com) (email contact: info@genomesystems.com) and

Research Genetics, Huntsville, Alabama (web address: www.resgen.com  
 ) (email contact: info@resgen.com) and  
 RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:  
 www.rzpd.de)

Seq primer: T3 ET from Amersham

High quality sequence stop: 434.

## FEATURES

## source

1..507  
 /organism="Dario rerio"  
 /db\_xref="taxon:7955"  
 /clone="4234174"  
 /clone\_lib="zebrafish gridded kidney"  
 /sex="mixed"  
 /tissue\_type="kidney pooled from 300 wild type adults"  
 /lab\_host="XL0LP"  
 /note="Organ: kidney; Vector: pBK-CMV; Site\_1: EcoRI;  
 Site\_2: XhoI; Oligo dt cDNA library constructed from mRNA  
 pooled from pooled kidney tissue from 300 adult  
 zebrafish."

BASE COUNT 108 a 150 c 107 g 141 t 1 others

## ORIGIN

## Alignment Scores:

Pred. No.: 2.6e-23 Length: 507  
 Score: 309.50 Matches: 59  
 Percent Similarity: 66.14% Conservative: 25  
 Best Local Similarity: 46.46% Mismatches: 39  
 Query Match: 15.99% Indels: 4  
 DB: 12 Gaps: 2

US-10-023-586B-2 (1-360) x BF938665 (1-507)

Qy 43 CysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe 62

Db 102 TGCAGCATCGACGAGTCCACAGTATGTCATCTACCGTTGCTGCTGCTTAACCTTC 161

Qy 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrp-----GlyGlnThr 80

Db 162 TTGCTCAGCCCTCGCTCACTCCACGCTCTCTCCGCTCTCCGATGCAAGAGTAGC 221

Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsn-LeuMetValAlaAspLeuLeuTy 100

Db 222 CGACAGTGGAAACACTTCGCTCATCTACATGGTCAATTCTGCCCTCTACGGATTTAATGTA 281  
 Qy 100 rValLeu---LeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpPropheG1 119  
 Db 282 CAGCTTCTCGCTTCGTTCCCTCGTGGCCAGCTACTTAATGCTGACCATTTGGGTTTCGG 341  
 Qy 119 yGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLe 139  
 Db 342 GGACTTCATGTGCAGGCTCGTTCGTTCTTTTCTTCAACCTGTACTGCAGCATCTT 401  
 Qy 139 uLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLe 159  
 Db 402 CTTCTTACCTGCATATCCGTCATCGGTACATGGGCATCTGCCATCCCAATCAGAACGAT 461  
 Qy 159 uProTyrArgThrArgArg 165  
 Db 462 CGCCCTGNAAAGCAAGCGG 480

Search completed: May 27, 2003, 00:31:18

Job time : 1784 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:45:35 ; Search time 37 Seconds  
(without alignments)  
1296.493 Million cell updates/sec

Title: US-10-023-586b-2  
Perfect score: 1936  
Sequence: 1 MSLILLPSGRSGSRGAL.....QNKLGHPAGRKRCPLNRS:360

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
A_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query		Description	
No.	Score	Match	Length	ID	
1	1798	92.9	333	22	AAE04125 Human G protein-co
2	1542	79.6	295	23	AAE11251 Human P2Y-like G p
3	1185	61.2	339	23	AAE18640 Human G-protein.co
4	921	47.6	170	22	AAE09335 Human nGPCR12... Ho
5	421	21.7	374	22	AAE04390 Turkey P2Y nucleot
6	416	21.5	537	23	AAU74538 Human P2Y purinoc
7	395	20.4	328	17	AAE01224 Mouse pancreas G-p
8	394.5	20.4	377	22	AAE01144 Human purinergic r
9	392.5	20.3	328	17	AAE01225 Human placenta G-p
10	387.5	20.0	328	22	AAE04393 Human P2-purinergi

11	387.5	20.0	377	22	AAE04392	Human P2-purinergi
12	387.5	20.0	377	22	AAE01143	Human purinergic r
13	371	19.2	373	23	AAE20604	Mus musculus GPCR
14	370.5	19.1	375	16	AAE72457	Human P20 receptor
15	370	19.1	97	22	AAU25585	Human G Protein-Co
16	368.5	19.0	328	18	AAW09433	Human placenta pur
17	367	19.0	365	18	AAW23606	Human P2Y4 recepto
18	367	19.0	365	22	AAE04391	Human P2Y4 pyrimid
19	329.5	17.0	259	21	AAE45375	Gene 37 human secr
20	329	17.0	373	23	AAU10985	Purinergic recepto
21	328	16.9	373	22	AAE04389	Human P2-purinergi
22	328	16.9	373	23	AAU10983	Purinergic recepto
23	328	16.9	373	23	AAU10984	Purinergic recepto
24	281.5	14.5	336	22	AAE09711	Human nGPCR54 #2.
25	278.5	14.4	337	22	AAU04375	Human G-protein co
26	278.5	14.4	337	22	AAU04584	Human G-protein co
27	278.5	14.4	337	23	AAU15399	Human G-protein co
28	278.5	14.4	337	23	ABH81902	Human G-protein co
29	278.5	14.4	337	23	ABH83819	Human P2Y-like rec
30	278.5	14.4	337	23	ABH81803	Human AXOR89 (G-pr
31	278.5	14.4	337	23	ABH79438	Human P2Y1-li. Ho
32	278.5	14.4	337	23	AAU77600	Human P2Y1-like G
33	278.5	14.4	337	23	AAO14027	Human purinergic-r
34	278.5	14.4	337	23	AAE16171	Human G-protein co
35	278	14.4	337	22	AAW79297	Human protein SEQ
36	273.5	14.1	344	17	AAW04247	Human G-protein co
37	273.5	14.1	344	21	AAE15413	Human G-protein co
38	273.5	14.1	344	21	AAE67356	Human G-protein co
39	264.5	13.7	370	23	ABP61510	Human NF-kB activa
40	258.5	13.4	370	19	AAW62597	Human 7-transmembr
41	258.5	13.4	370	23	ABP61511	Human NF-kB activa
42	250.5	12.9	374	19	AAW51406	Human NF-kB activa
43	249	12.9	339	15	AAE53752	Seven transmembran
44	249	12.9	339	18	AAE07617	Human G-protein th
45	249	12.9	339	19	AAW48733	Human R12 seven tr

ALIGNMENTS

RESULT 1	
AAE04125	AAE04125 standard; Protein; 333 AA.
ID	AAE04125 standard; Protein; 333 AA.
XX	
AC	AAE04125;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human G protein-coupled receptor GPR71.
XX	
KW	Human: guanosine triphosphate binding protein-coupled receptor;
KW	G protein-coupled receptor; GPR8; GPR12; GPR16; GPR21; GPR40;
KW	GPR47; GPR51; GPR71; GPR72; GPR72; GPR72; GPR72; GPR72;
KW	Alzheimer's disease; cytostatic; hepatotropic; nontropic;
KW	neuroprotective; gene therapy; peptide therapy.
OS	Homo sapiens.
XX	
PN	WO200148188-A1.
XX	
PD	05-JUL-2001.
XX	
PF	28-DEC-2000; 2000WO-JP09408.
XX	
PR	28-DEC-1999; 99JP-0375152.
XX	
PA	31-MAR-2000; 2000JP-0101339.
XX	
XX	(HELI-) HELIX RES INST.
PI	Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
PI	Sugiyama T, Kishimoto T, Kanzaki K, Yasuda S, Inoue Y;
XX	
DR	WPI; 2001-425662/45.

DR N-PSDB; AAH73516.  
 XX New DNA encoding guanosine triphosphate binding protein coupled  
 XX receptors and their expression products for screening potential  
 PT anticancer and neurotropic drugs and in diagnosis of these diseases  
 XX  
 XX Example 1: Page 132-135; 170pp; Japanese.  
 XX  
 XX The invention relates to nine human guanosine triphosphate binding  
 CC protein (G protein)-coupled receptors designated GPRv8, GPRv12, GPRv16,  
 CC GPRv21, GPRv40, GPRv47, GPRv51, GPRv71 and GPRv72, and to the  
 CC genes encoding them. These genes and proteins and antibodies against  
 CC the protein are useful in the treatment, prevention, diagnosis and  
 CC investigation of diseases associated with G protein-coupled receptors,  
 CC including cancer, cirrhosis of the liver and Alzheimer's disease.  
 CC The present sequence is a G protein-coupled receptor of the invention.  
 XX  
 XX Sequence 333 AA;

Query Match 92.9%; Score 1798; DB 22; Length 333;  
 Best Local Similarity 99.7%; Pred. No. 3.8e-191;  
 Matches 331; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 29 MEKVDNMTSQEGLCQFSEKQVYLSLAYSLIIFILGLPLNGTVLWHSGQTKRWSCATT 88  
 1 MEKVDNMTSQEGLCQFSEKQVYLSLAYSLIIFILGLPLNGTVLWHSGQTKRWSCATT 60  
 QY 89 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 148  
 61 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 120  
 QY 149 FLGVCHPLCSPYRTRHAWLGTSTTVALVQLLPTLAFSHTDYINGOMIYDWTSGEN 208  
 121 FLGVCHPLCSPYRTRHAWLGTSTTVALVQLLPTLAFSHTDYINGOMIYDWTSGEN 180  
 QY 209 FDRLFAYGIVLTLSGFLSLGHFGVLTGDPQDQARCEPHEDRQHSQVHPDPTGVM 268  
 181 FDRLFAYGIVLTLSGFLSLGHFGVLTGDPQDQARCEPHEDRQHSQVHPDPTGVM 240  
 QY 269 PLHPLFCALPYHSLLLPHHLLSAFSGLPALDQSGCGGLQDMASGECEQ 328  
 241 PLHPLFCALPYHSLLLPHHLLSAFSGLPALDQSGCGGLQDMASGECEQ 300  
 QY 329 GKNRVRLQLKRONKLGEPHAGRCPCGLNRS 360  
 301 GKNRVRLQLKRONKLGEPHAGRCPCGLNRS 332

RESULT 2  
 AAU11251  
 ID AAU11251 standard; Protein; 295 AA.

XX AAU11251;

XX 26-FEB-2002 (first entry)

XX Human P2Y-like G protein-coupled receptor.

DE Human; P2Y-like G protein-coupled receptor; GPCR; COPD;  
 XX chronic obstructive pulmonary disease; nervous system disease;  
 KW Parkinson's disease; multiple sclerosis; dementia; stroke;  
 KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;  
 KW bacterial infection; fungal infection; protozoan infection;  
 KW viral infection; pain; cancer; anorexia; bulimia; asthma;  
 KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;  
 KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;  
 KW psychotic disorder; neurological disorder; anxiety; schizophrenia;  
 KW manic depression; delirium; severe mental retardation; dyskinesia.

XX Homo sapiens.

OS W0200185764-A2.

XX

PD 15-NOV-2001.  
 XX  
 XX 09-MAY-2001; 2001WO-EP05244.  
 XX  
 XX 11-MAY-2000; 2000US-203582P.  
 PR 21-FEB-2001; 2001US-269857P.  
 XX  
 XX (FARB ) BAYER AG.  
 PA  
 XX Ramakrishnan S;  
 PI  
 XX WPI; 2002-075242/10.  
 DR N-PSDB; AAS17746.

New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders

Claim 25; Fig 2; 114pp; English.

The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR. In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, osteoporosis, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies, psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic depression, delirium, severe mental retardation or dyskinesias). The present sequence is the P2Y-like GPCR of the invention.

Sequence 295 AA;

Query Match 79.6%; Score 1542; DB 23; Length 295;  
 Best Local Similarity 98.6%; Pred. No. 9.6e-163;  
 Matches 284; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 29 MEKVDNMTSQEGLCQFSEKQVYLSLAYSLIIFILGLPLNGTVLWHSGQTKRWSCATT 88  
 1 MEKVDNMTSQEGLCQFSEKQVYLSLAYSLIIFILGLPLNGTVLWHSGQTKRWSCATT 60  
 QY 89 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 148  
 61 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 120  
 QY 149 FLGVCHPLCSPYRTRHAWLGTSTTVALVQLLPTLAFSHTDYINGOMIYDWTSGEN 208  
 121 FLGVCHPLCSPYRTRHAWLGTSTTVALVQLLPTLAFSHTDYINGOMIYDWTSGEN 180  
 QY 209 FDRLFAYGIVLTLSGFLSLGHFGVLTGDPQDQARCEPHEDRQHSQVHPDPTGVM 268  
 181 FDRLFAYGIVLTLSGFLSLGHFGVLTGDPQDQARCEPHEDRQHSQVHPDPTGVM 240  
 QY 269 PLHPLFCALPYHSLLLPHHLLSAFSGLPALDQSGCGGLQDMASGECEQ 316  
 241 PLHPLFCALPYHSLLLPHHLLSAFSGLPALDQSGCGGLQDMASGECEQ 288

RESULT 3

AAE18640  
ID AAE18640 standard; Protein; 339 AA.  
XX  
AC AAE18640;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human G-protein coupled receptor (GCREC-1).  
XX  
KW Human: G-protein coupled receptor; GCREC-1; cell proliferative disorder;  
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;  
KW neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;  
KW virucide; receptor.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 89..109  
FT /note="Transmembrane domain"  
FT Domain 130..149  
FT /note="Transmembrane domain"  
XX  
XX W0200210387-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 25-JUL-2001; 2001WO-US23433.  
XX  
XX 27-JUL-2000; 2000US-221478P.  
XX 23-AUG-2000; 2000US-223268P.  
XX 21-AUG-2000; 2000US-227054P.  
XX 08-SEP-2000; 2000US-231121P.  
XX 13-SEP-2000; 2000US-232243P.  
XX 15-SEP-2000; 2000US-232691P.  
XX 22-SEP-2000; 2000US-235146P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
XX Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walia NK;  
XX Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L;  
XX Graul RC, Warren BA, Lee EA, Ding L;  
XX  
XX WPI; 2002-188744/24.  
XX N-PSDB; AAD29667.  
XX  
XX New human G-protein coupled receptor polypeptide for diagnosis,  
XX prevention and treatment of cell proliferative, neurological,  
XX cardiovascular, gastrointestinal, autoimmune/inflammatory, and  
XX metabolic disorders  
XX  
XX Claim 1; Page 114-115; 150pp; English.  
XX  
XX The invention relates to novel human G-protein coupled receptors  
XX (GCREC) and their encoding polynucleotides. GCREC is useful as an  
XX immunogen for preparing monoclonal and polyclonal antibodies. GCREC is  
XX useful for diagnosing, treating and preventing a cell proliferative  
XX disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder  
XX (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's  
XX disease, Parkinson's disease), a cardiovascular disorder (e.g.,  
XX atherosclerosis, hypertension, myocardial infarction), gastrointestinal  
XX disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/  
XX inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS),  
XX allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder  
XX (e.g., diabetes, obesity, osteoporosis), and viral infections. GCREC is  
XX useful in a number of drug screening techniques, and to analyse the  
XX proteome of a tissue or cell type. GCREC is useful for creating knockin

CC humanised animals or transgenic animals to model human diseases, in  
CC somatic or germline gene therapy, to generate a transcript image of a  
CC tissue or cell type, for detecting differences in the chromosomal  
CC location due to translocation, inversion, etc., among normal, carrier  
CC or affected individuals, and as hybridization probes for mapping  
CC naturally occurring genomic sequences. GCREC is useful in Southern or  
CC northern analysis, dot blot or other membrane-based technologies, in PCR  
CC technologies, in dipstick, pin, multiformat enzyme linked immunosorbent  
CC (ELISA)-like assays, and in microarrays utilising fluids or tissues from  
CC patients to detect altered GCREC expression. The present sequence is  
CC human GCREC-1.  
XX  
SQ Sequence 339 AA;  
Query Match 61.2%; Score 1185; DB 23; Length 339;  
Best Local Similarity 99.6%; Pred. No. 6.1e-123;  
Matches 225; Conservative 0; Mismatches 1; Indels: 0; Gaps 0;  
QY 1 MLSILLPSRSGSRGRRGALLLEGASRDMKXDMNTSQEQLGCFSEKVKQVLSLAYSI 60  
DB 1 MLSILLPSRSGSRGRRGALLLEGASRDMKXDMNTSQEQLGCFSEKVKQVLSLAYSI 60  
QY 61 IFILGLPLNGTVLWHSWGQTKRSCATYLVNLMVADLLYVLLPFLIITYSLDDRWPFGE 120  
DB 61 IFILGLPLNGTVLWHSWGQTKRSCATYLVNLMVADLLYVLLPFLIITYSLDDRWPFGE 120  
QY 121 LKCLVHFLFYINLYGSILLTLCISVHQFLGVCPLCSLPYTRRHAWLGTSTTWALVVL 180  
DB 121 LKCLVHFLFYINLYGSILLTLCISVHQFLGVCPLCSLPYTRRHAWLGTSTTWALVVL 180  
QY 181 QLLPTLAFSHPTDYINGQMIWYDMTSQENFDRLFAYGIVLTLSGFLS 226  
DB 181 QLLPTLAFSHPTDYINGQMIWYDMTSQENFDRLFAYGIVLTLSGFLS 226  
RESULT 4  
AAG80935  
ID AAG80935 standard; Protein; 170 AA.  
XX  
AC AAG80935;  
XX  
DT 28-AUG-2001 (first entry)  
XX  
DE Human nGPCR12.  
XX  
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;  
KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
KW cardiovascular disease; proliferative disorder; hormonal disorder;  
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
KW attention deficit-hyperactivity disorder/attention deficit disorder;  
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
PN W0200136473-A2.  
XX  
PD 25-MAY-2001.  
XX  
XX 16-NOV-2000; 2000WO-US31581.  
XX  
XX 16-NOV-1999; 99US-0165838.  
XX 17-NOV-1999; 99US-0166071.  
XX 19-NOV-1999; 99US-0166678.  
XX 22-DEC-1999; 99US-0173396.  
XX 28-FEB-2000; 2000US-0184129.  
XX 28-FEB-2000; 2000US-0185421.  
XX 28-FEB-2000; 2000US-0185554.  
XX 02-MAR-2000; 2000US-0186530.  
XX 03-MAR-2000; 2000US-0186811.  
XX 09-MAR-2000; 2000US-0188114.







CC used to screen agonists and antagonists that modulate G-PCR  
 CC activity, to raise antibodies and to develop assay systems.

SQ Sequence 328 AA;

Query Match 20.4%; Score 395; DB 17; Length 328;  
 Best Local Similarity 31.9%; Pred. No. 4.4e-35;  
 Matches 106; Conservative 55; Mismatches 137; Indels 34; Gaps 9;

Qy 29 MEKVDNMTSQEGL----COFSEKQKVYLSLAYSIIFILGLPLNGTVMHWSGOTKRWK 84  
 Db 1 MEQ-DNGTIQAPLPPTTCVYREDFKRLLTPVYSVILVGVGLPNICVIAQCASRRRLT 59  
 Qy 85 CATTYLVNMLVADLYVL-LPFLIITYSDDRWPFGEGLCKLVHFLFYINLYGSILLTLC 143  
 Db 60 RSAVYTNLALADLMYACSLPLIINYARGDHWPFGLACRFVRFYANLHGSILFLTC 119  
 Qy 144 ISVHQFLGVCHPLCSLPR-TRRHAWLGTSTWALVVLQLLPTLAFSTHDYINGOMIWD 202  
 Db 120 ISFYRLGICHPLASWHRGRGGRRAAVVGVVWLVATAQAQLPTAVFAATGIQRNRTCYD 179  
 Qy 203 MTSQENFORLAYGIVLTLSGFL----SLHGFGVLETDGQEPDQARGEHEDROHSPSQ 258  
 Db 180 LSPPTLSRYLPYGMALIVIGFLPFIALLACYCRMARRLCRDGPAGPVAQERSKAAR 239  
 Qy 259 VPHDPTGVNPLHPLFCALPYHSLLLPHLL-SAFSGPLALDGSQCGLQDMEASGECEQL 317  
 Db 240 M-----AVVAFAVFAISLPHITKTAVLAVRSTPGVSCPVLTFFAAYKQTR 287  
 Qy 318 PQPS-----PVLSF-----KGGKNVRLLQKL 339  
 Db 288 PFASVNSVLDPIFYFTOOKFRROPHDLLQRL 319

RESULT 8

ID AAE01144 standard; Protein; 377 AA.

XX AAE01144;

DT 17-JUL-2001 (first entry)

XX Human purinergic receptor P2Y2 protein.

XX Human; purinergic receptor; P2Y2; antidepressant; vulnery; hypotensive;  
 KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
 KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
 KW irritable bowel disorder; reproductive system disorder; hypertension;  
 KW peripheral vascular disease; immune system disorder; chronic bronchitis;  
 KW premature ejaculation; asthma; neuromuscular disease.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 258 /note= "Encoded by TFC"

FT Misc-difference 334 /note= "Wild type Arg substituted with Cys"

FT Misc-difference 350 /note= "Wild type Gly substituted with Glu"

FT Misc-difference 359 /note= "Wild type Phe substituted with Ser"

XX US6214581-B1.

XX 10-APR-2001.

XX 13-NOV-1998; 98US-0191136.

XX 16-JAN-1998; 98US-0071298.

PR 16-JAN-1998; 98US-0071669.

PR 16-JAN-1998; 98US-0008185.

PR 16-JAN-1998; 98US-0008526.

XX

PA (ABBO ) ABBOTT LAB.

PI Lynch KJ, Burgard EC, Van Biesen T;

XX WPI; 2001-315459/33.

DR N-PSDB; AAD04981.

XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor  
 PT polypeptide useful for identifying potentially therapeutic compounds  
 PT that modulate or otherwise interact with P2X containing receptors -

PS Example 14; Fig 12B; 53pp; English.

XX The present sequence is human P2Y2 receptor protein. P2X  
 CC receptors are ligand-gated ion channels while P2Y receptors operate  
 CC generally through a G-protein coupled system. P2X purinoreceptor drugs  
 CC are potential therapeutic agents in several disorders including central  
 CC nervous system or peripheral nervous system conditions, e.g., epilepsy,  
 CC pain, depression, neurodegenerative disorders, disorders of the skeletal  
 CC muscle such as neuromuscular diseases, disorders of the reproductive  
 CC system, asthma, peripheral vascular disease, hypertension, immune system  
 CC disorders, irritable bowel disorder, premature ejaculation, cystic  
 CC fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity  
 CC of extracellular nucleotide triphosphates to regulate chloride secretion  
 CC in human airway epithelia.

XX Sequence 377 AA;

Query Match 20.4%; Score 394.5; DB 22; Length 377;

Best Local Similarity 32.8%; Pred. No. 6.1e-35;

Matches 117; Conservative 40; Mismatches 121; Indels 79; Gaps 12;

Qy 43 COFSEKQKVYLSLAYSIIFILGLPLNGTVMHWSGOTKRWKSCATTYLVNMLVADLYVL 102  
 Db 25 CRNEDFKYVLLPVSYGVVGLCLNAVALYIFLCRLKTNASTYWFHLAYSALYAA 84  
 Qy 103 -LPFLIITYSDDRWPFGEGLCKLVHFLFYINLYGSILLTLCISVHQFLGVCHPLCSLPY 161  
 Db 85 SLPLLVYVYARGDHWPFSTVLCKLVRFELTYNLYSILFLTCISVHRCGLVRLRLRW 144  
 Qy 162 RTRRHAWLGTSTWALVVLQLLPTLAFSTHDYINQMTWYDTSOENFDRLFAYGIVLT 221  
 Db 145 GRARTYRVAAGAVVVLACQAPVLYFTTSARGGRVCHDTSAPELFSRFVAYSSVM-- 202  
 Qy 222 SGFSLLLGHFGVLF-----TDGQEPDQARGEHEDROHSPSQVHP 261  
 Db 203 ---LGLL--FAVPPAVILVCYVLMARRLLKPAYGTSG-----GLPRAKRKSVRT--- 246  
 Qy 262 DHPTGVNPLHPLF--CALPYH-----SLLPHLLLSAFS-----GLPALDGSQCG 304  
 Db 247 -----IAVVLAVFALCFLPFHVTRTYYSFRSLDLSCHTLNAINMAYKVTRPLASNSC- 300  
 Qy 305 LQDMEASGECEQLPQSPVLSFKGKNVRLLQKLQKNKLGHPAGRKRCP-GLNRS 360  
 Db 301 -----LDPVLYFLAGQLRVRFARDAKP--PTGSPATPARCRLGLRRS 341

RESULT 9

AAR91225

ID AAR91225 standard; Protein; 328 AA.

XX AAR91225;

DT 26-AUG-1996 (first entry)

XX Human placenta G-protein coupled receptor protein.

XX G-protein coupled receptor protein; G-PCR; agonist; antagonist;  
 KW cystic fibrosis; incontinence; diabetes; diagnosis; therapy.

XX Homo sapiens.

PN WO9605302-A1.  
 XX 22-FEB-1996.  
 XX 10-AUG-1995; 95WO-JP01599.  
 XX 11-AUG-1994; 94JP-0189272.  
 PR 11-AUG-1994; 94JP-0189273.  
 PR 11-AUG-1994; 94JP-0189274.  
 PR 30-SEP-1994; 94JP-0236356.  
 PR 30-SEP-1994; 94JP-0236357.  
 PR 28-NOV-1994; 94JP-0270017.  
 PR 28-DEC-1994; 94JP-0326611.  
 PR 20-JAN-1995; 95JP-0007177.  
 PR 16-MAR-1995; 95JP-0057186.  
 PR 19-APR-1995; 95JP-0093989.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA Fujii R, Fukusumi S, Hinuma S, Hosoya M, Ohgi K;  
 PI Ohtaki T;  
 XX WPI; 1996-139698/14.  
 DR N-PSDB; AAT18368.  
 XX G-protein coupled receptor protein DNA and protein - also methods  
 PT for isolating (ant)agonists for treatment of cystic fibrosis,  
 PT incontinence and diabetes  
 XX Claim 6; Page 267-68; 360pp; English.  
 XX A novel human placenta-derived G-protein coupled receptor protein  
 CC (G-PCR) (AAR91225) was identified as the product of cDNA clone  
 CC pHA2-17 (AAT18368). The protein can be obtd. by expression of the  
 CC cDNA clone in transformed host cells. It was classified as a  
 CC purinoceptor. G-PCRs (see also AAR91217-24 and AAR91227-33) can be  
 CC used to screen agonists and antagonists that modulate G-PCR  
 CC activity, to raise antibodies and to develop assay systems.  
 XX Sequence 328 AA;  
 XX Query Match 20.3%; Score 392.5; DB 17; Length 328;  
 XX Best Local Similarity 33.0%; Pred. No. 8.4e-35;  
 XX Matches 109; Conservative 42; Mismatches 142; Indels 37; Gaps 9;  
 QY 33 DMNTSQEQL-----CQSEKVKQVLSLAYSIIFILGPLNGTVLHWSGQTKRWSATT 88  
 DB 4 DNGTGQALGPLPTTCVYRENEFKQLLPVYSAVLAAGLPNLCVITQICTSRRAITAV 63  
 QY 89 YLVNLMVADLLYL-LPFLITYSLDRWPFGECLLKVHFLFYINLYGSLILLTCISVH 147  
 DB 64 YTLNLADLLYACSLPPLLYNYAAGDHPGDFACRLVFLFVANLHGSILFLTCISFQ 123  
 QY 148 QFLGVCHPLCSLYR-TRRIHNLGTSTWALVQLLPTLAFSHDYNQOMIYDWTQS 206  
 DB 124 RYLIGCHPLAPWKKRGRRRAALVCVTVLAVTTQCLPTAIFATGTRNRTVCYDLSP 183  
 QY 207 ENFDRLFAYGIVLTLSGL-----SLGHFGVLTGDEPQARCEPHEDRQSPSQVHPD 262  
 DB 184 ALATHYMPYGMATLVIGFLFPFALLACACYLAC-----RLCRDGDGPAEPAVQ 231  
 QY 263 HPTGVNPLHPLFCALPVLHSLHLL-SAFSGLPALDGGSCGLQDMASGECEQLPOP- 320  
 DB 232 ERGKAARMVAVVNAFAISFLPHITKXTAYLAVGSTPGVPCIV--LEAFAAYKGRTRF 289  
 QY 321 -----SPVLSE-----KGGKNRVLQKL 339  
 DB 290 ASANSVLDPLLFYFTQKKFRFRPHELIQKL 319  
 RESULT 10  
 AAE04393  
 ID AAE04393 standard; Protein; 328 AA.

XX AAE04393;  
 XX 04-SEP-2001 (first entry)  
 XX Human P2-purinergic receptor subtype, P2Y6.  
 XX Human; P2-purinergic receptor; P2Y6; cardiant; vasotropic; thrombolytic;  
 KW cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina;  
 KW myocardial infarction; ischaemic attack; preeclampsia; bleeding disorder;  
 KW carotid endarterectomy; vascular graft surgery; brain disorder; migraine;  
 KW vascular injury; schizophrenia; eating disorder; depression; angiodiastly;  
 KW peripheral vascular disease; platelet aggregation; restenotic; embolism;  
 KW thrombocytopaenic purpura; stroke; pertussis toxin-sensitive G protein;  
 KW GI; disseminated intravascular coagulation; thrombosis.  
 XX Homo sapiens.  
 OS WO200146454-A1.  
 XX 28-JUN-2001.  
 XX 26-DEC-2000; 2000WO-US34998.  
 XX 23-DEC-1999; 99US-0171622.  
 XX (CORT-) COR THERAPEUTICS INC.  
 XX Conley PB, Jantzen H, Ramakrishnan-Dubridge V, Julius DJ;  
 PI Holloper G;  
 XX WPI; 2001-418082/44.  
 XX Novel isolated ADP receptor, termed P2Y12 receptor polypeptide, useful  
 PT for identifying binding partners and for diagnostic applications -  
 XX Disclosure; Page 99-101; 108pp; English.  
 XX The invention relates to ADP (adenosine 5'-diphosphate) receptor, termed  
 CC as P2Y12 receptor and its corresponding cDNA molecule. P2Y12 receptor is  
 CC the subtype of P2-purinergic receptor. The P2Y12 receptor is expressed  
 CC selectively in the platelets and brain, and couples to a pertussis toxin-  
 CC sensitive G protein (Gi). P2Y12 receptor is a G protein-coupled receptor  
 CC that responds to ADP. The invention also relates to a method for  
 CC identifying an agent which is useful for modulating acute myocardial  
 CC infarction, unstable angina, chronic stable angina, transient ischaemic  
 CC attacks, strokes, peripheral vascular disease, preeclampsia, deep venous  
 CC thrombosis, embolism, disseminated intravascular coagulation, thrombotic  
 CC thrombocytopaenic purpura or a bleeding disorder; thrombotic and  
 CC restenotic complications following angioplasty, carotid endarterectomy,  
 CC post CABG (coronary artery bypass graft) surgery, vascular graft surgery,  
 CC stent placements or insertion of endovascular devices and prostheses.  
 CC P2Y12 receptor is useful for identifying binding partners and for  
 CC diagnostic applications. P2Y12 receptor provides targets for screening  
 CC synthetic small molecules and combinatorial or naturally occurring  
 CC compound libraries to regulate platelet aggregation, vascular injury, or  
 CC disease as well as schizophrenia, eating disorders, depression, migraine  
 CC and other brain disorders. The present sequence is a subtype of human  
 CC P2-purinergic receptor subtype, P2Y6 related to the invention.  
 XX Sequence 328 AA;  
 XX Query Match 20.0%; Score 387.5; DB 22; Length 328;  
 XX Best Local Similarity 32.7%; Pred. No. 3e-34;  
 XX Matches 108; Conservative 42; Mismatches 143; Indels 37; Gaps 9;  
 QY 33 DMNTSQEQL-----CQSEKVKQVLSLAYSIIFILGPLNGTVLHWSGQTKRWSATT 88  
 DB 4 DNGTGQALGPLPTTCVYRENEFKQLLPVYSAVLAAGLPNLCVITQICTSRRAITAV 63  
 QY 89 YLVNLMVADLLYL-LPFLITYSLDRWPFGECLLKVHFLFYINLYGSLILLTCISVH 147  
 DB 64 YTLNLADLLYACSLPPLLYNYAAGDHPGDFACRLVFLFVANLHGSILFLTCISFQ 123





Db 85 SLPLLVYYARGDHPFSTVLCVLRFLFYNTLYCSILFLTCISVHRCLGVLRLPLHSLRW 144  
 QY 162 RTRRHAWLGSTTTWALVVLQLLPTLAFSHTDYINGOMIWDMTSOENFDRLFAYGIVLTL 221  
 Db 145 GRARVARRAAVWVVLVACQAPVLYFVTTSVRGFRITCHDSARELFSHFVAYSSVM-- 202  
 QY 222 SGFLSLLGHFGVLF-----TDGQEPDQARGEHPEDRO--HSPSOV 259  
 Db 203 ---LGLL--FAVPFVILVCVLMARLLKPAYGTG-----GLPRAKRKSVRTIALV 250  
 QY 260 HPDHTPTGVWPLHPLCALPYH-----SLLPHHLLSAFS-----GLPALDGSOCG 304  
 Db 251 -----LAVFAL-----CFLPFHVTRTLYSFRSLDLSCHTLNAINMAYKITRPLASANS- 300  
 QY 305 LODMEASGECEQLPOPSVLFKGGKRNVRLLQKLR-----ONKLGEPHAPGR 351  
 Db 301 -----LDPVLYLAGORLVRFARDAKPTEPTPSQARRKLGHLHPNR 343

RESULT 14  
 AAR72457  
 ID AAR72457 standard; Protein; 375 AA.  
 XX  
 AC AAR72457;

29-NOV-1995 (first entry)

DE Human P20 receptor.  
 XX  
 KW Epithelial mucosa; mucus; cystic fibrosis; asthma;  
 KW chronic bronchitis.

OS Homo sapiens.

XX WO9510538-A.

XX PD 20-APR-1995.

XX PF 04-OCT-1994; 94WO-US11260.

XX PR 15-OCT-1993; 93US-0138137.

XX PA (UMOR ) UNIV MISSOURI.

XX PA (UTNC-) UNIV NORTH CAROLINA.

XX PI Boucher RC, Erb LJ, Harden TK, Lustig KD, Parr CE;

XX PI Sullivan DM, Turner JT, Weisman GA;

DR WPI; 1995-169967/22.

DR N-PSDB; AAQ88134.

XX DNA encoding human P20 receptor and null cells expressing the  
 XX receptors - for stimulating or inhibiting growth of cultures of  
 XX mammalian cells, and for treating diseases of airway epithelial  
 XX mucosa, e.g. asthma.

XX PS Disclosure; Page 34; 47pp; English.

XX The sequence is that of the human P20 receptor. The receptor may  
 CC be used to treat diseases of epithelial mucosal surfaces by  
 CC enhancing or inhibiting mucus prodn. It may be used to treat  
 CC cystic fibrosis, asthma and chronic bronchitis.

XX SQ Sequence 375 AA;

Query Match 19.1%; Score 370.5; DB 16; Length 375;

Best Local Similarity 32.7%; Pred. No. 2.8e-32;

Matches 115; Conservative 42; Mismatches 124; Indels 71; Gaps 13;

QY 43 CQFSEKQVLSLAYSIIFILGLPLNGTVLWHSWGQKRWSCATTYLVNLMVADLLXVL 102

Db 25 CRFNEEDLAVLLPVSGVGVVCLGICLNAVGLYIFLCRLKTNASTYMFHLAVSDALYAA 84

QY 103 -LPFLIITYSLDDRWPFGBELCKLVHFLFYINLYSGIILLTCTISVHQFLGVCHPLCSLPY 161  
 Db 85 SLPLLVYYARGDHPFSTVLCVLRFLFYNTLYCSILFLTCISVHRCLGVLRLPLHSLRW 144  
 QY 162 RTRRHAWLGSTTTWALVVLQLLPTLAFSHTDYINGOMIWDMTSOENFDRLFAYGIVLTL 221  
 Db 145 GRARVARRAAVWVVLVACQAPVLYFVTTSVRGFRITCHDSARELFSHFVAYSSVM-- 201  
 QY 222 SGFLSLLGHFGVLF-----TDGQEPDQARGEHPEDROHSPSOVHP 261  
 Db 202 ---LGLL--FAVPFVILVCVLMARLLKPAYGTG-----GLPRAKRKSVRT--- 245  
 QY 262 DHPDHTPTGVWPLHPLCALPYH-----SLLPHHLLSAFSGLPALDGSOCGLQDME 309  
 Db 246 -----LAVFAL-----CFLPFHVTRTLYSFRSLDLSCHTLNAIN-----MAYKVTRLA 293  
 QY 310 ASGECEQLPOPSVLFKGGKRNVRLLQKLRQKLGHEPA-GRKRCPCGLNRS 360  
 Db 294 SANS-----LDPVLYLAGORLVRFARDAKP-PTGSPATPARRTLGLRRS 339

RESULT 15

AAU25585

ID AAU25585 standard; Protein; 97 AA.

XX AC AAU25585;

XX DT 18-DEC-2001 (first entry)

XX DE Human G Protein-Coupled Receptor (GPCR) polypeptide #32.

XX KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;  
 KW attention deficit disorder; anxiety; depression; bipolar disorder;  
 KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;  
 KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;  
 KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;  
 KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;  
 KW viral infection; antineoplastic; neuroleptic; nootropic; tranquiliser;  
 KW antidepressant; anorectic; gene therapy.

XX OS Homo sapiens.

XX PN WO200162797-A2.

XX PD 30-AUG-2001.

XX PF 23-FEB-2001; 2001WO-US05676.

XX PR 23-FEB-2000; 2000US-0184247.

XX PR 23-FEB-2000; 2000US-0184303.

XX PR 23-FEB-2000; 2000US-0184304.

XX PR 23-FEB-2000; 2000US-0184305.

XX PR 02-MAR-2000; 2000US-0186457.

XX PR 03-MAR-2000; 2000US-0186810.

XX PR 09-MAR-2000; 2000US-0188064.

XX PR 13-MAR-2000; 2000US-0188880.

XX PR 03-APR-2000; 2000US-0194344.

XX PR 23-JUN-2000; 2000US-0213861.

XX PR 11-JUL-2000; 2000US-0217369.

XX PR 11-JUL-2000; 2000US-0217370.

XX PR 14-JUL-2000; 2000US-0218337.

XX PR 20-JUL-2000; 2000US-0218492.

XX PA (PHAA ) PHARMACIA & UPJOHN CO.

XX PI Vogeli G, Wood LS, Parodi LA, Lind P;

XX DR WPI; 2001-570628/64.

XX DR N-PSDB; AAS42837.

XX PT New isolated nucleic acid encoding a new G-protein coupled receptor  
 PT polypeptide for detecting receptor modulators that can treat mental

PT disorders, such as schizophrenia, anxiety, depression, or obesity ;  
XX  
PS Claim 35; Page 82; 279pp; English.  
XX  
CC Sequences AAU25554-AAU25516 represent human G-protein coupled receptor  
CC (GPCR) polypeptides of the invention. The proteins and their associated  
CC DNA sequences can be used to identify compounds which bind to GPCR  
CC polypeptides and in screening for compounds that modulate GPCR activity.  
CC By screening a human subject for the presence of mutations in GPCR DNA, a  
CC GPCR-related disorder or a genetic predisposition can be diagnosed. The  
CC sequences can also be used for treatment and prevention of mental  
CC disorders such as schizophrenia, attention deficit disorder, anxiety,  
CC depression, dementia and bipolar disorder, neurological disorders such as  
CC Huntington's disease, Parkinson's disease and Tourette's syndrome,  
CC metabolic disorders such as obesity, anorexia and type 2 diabetes,  
CC cardiovascular disorders such as thrombosis, myocardial infarction,  
CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and  
CC cancers.

XX SQ Sequence 97 AA;

Query Match 19.1%; Score 370; DB 22; Length 97;  
Best Local Similarity 100.0%; Pred. No. 4.9e-33;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSLTLLPSRSGSRGRRGALLLEGASRDMEKVDNMTSQEGLCQFSEKYKQVYLSLAYS I 60  
DB 24 MSLTLLPSRSGSRGRRGALLLEGASRDMEKVDNMTSQEGLCQFSEKYKQVYLSLAYS I 83  
QY 61 IFILGLPLNGTVLW 74  
DB 84 IFILGLPLNGTVLW 97

Search completed: May 22, 2003, 12:49:44  
Job time : 39 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:49:05 ; Search time 16 Seconds  
(without alignments)  
662.016 Million cell updates/sec

Title: US-10-023-586B-2  
Perfect score: 1936  
Sequence: 1 MSLTLLPSGRSGSRGAL.....QNKLGHPAGRKRCPLNRS 360

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	395	20.4	328	3	US-08-513-974B-39
2	395	20.4	328	3	US-08-513-974B-371
3	392.5	20.3	328	3	US-08-513-974B-56
4	392.5	20.3	328	3	US-08-513-974B-380
5	383.5	19.8	327	3	US-08-513-974B-372
6	374.5	19.3	375	1	US-08-442-134A-2
7	374.5	19.3	375	1	US-08-444-581B-2
8	374.5	19.3	375	1	US-08-446-088A-2
9	374.5	19.3	375	2	US-08-559-524A-3
10	374.5	19.3	375	2	US-08-749-707-3
11	369	19.1	373	3	US-08-513-974B-373
12	368.5	19.0	328	3	US-08-459-046-2
13	340	17.6	362	3	US-08-513-974B-374
14	326.5	16.9	373	2	US-08-559-524A-4
15	326.5	16.9	373	3	US-08-749-707-4
16	274.5	14.2	302	3	US-08-467-948A-30
17	274.5	14.2	302	3	US-08-467-947A-30
18	273.5	14.1	344	3	US-08-467-948A-8
19	273.5	14.1	344	3	US-08-467-947A-8
20	258.5	13.4	370	3	US-08-781-250-2
21	250.5	12.9	408	2	US-08-742-440A-6
22	249	12.9	339	1	US-08-153-848-44
23	249	12.9	339	2	US-08-812-871-3
24	249	12.9	339	3	US-09-299-843A-44
25	249	12.9	339	4	US-09-088-337B-44
26	249	12.9	339	5	PCT-US93-11153-44
27	249	12.9	339	5	PCT-US95-07180-2

28	248.5	12.8	391	1	US-07-816-283-2	Sequence 2, Appli
29	248.5	12.8	391	1	US-08-417-103-2	Sequence 2, Appli
30	248.5	12.8	391	1	US-08-417-103-14	Sequence 14, Appl
31	246	12.7	347	4	US-08-405-271A-24	Sequence 24, Appl
32	246	12.7	395	1	US-08-097-938-2	Sequence 2, Appli
33	246	12.7	395	1	US-08-097-938-5	Sequence 5, Appli
34	246	12.7	395	1	US-08-476-000-2	Sequence 2, Appli
35	246	12.7	395	1	US-08-476-000-5	Sequence 5, Appli
36	246	12.7	395	1	US-08-472-840-2	Sequence 2, Appli
37	246	12.7	395	1	US-08-472-840-5	Sequence 5, Appli
38	246	12.7	395	2	US-08-476-976-2	Sequence 2, Appli
39	246	12.7	395	2	US-08-476-976-5	Sequence 5, Appli
40	246	12.7	395	3	US-08-474-410-2	Sequence 2, Appli
41	246	12.7	395	3	US-08-474-410-5	Sequence 5, Appli
42	246	12.7	395	4	US-08-486-673B-2	Sequence 2, Appli
43	246	12.7	395	4	US-08-486-673B-5	Sequence 5, Appli
44	246	12.7	399	1	US-08-476-000-61	Sequence 61, Appl
45	246	12.7	399	1	US-08-472-840-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1  
US-08-513-974B-39  
; Sequence 39, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357



Db 60 RSAYVTNLNADLADLVACSLPLIYNYARGDHWPFGLACRFVRFYANLHGSILFLTC 119  
QY 144 ISVHQFLGVCHPLCSLPYR-TRRHAMLGSTTWALVVLQLLPTLAFSHDYINGOMIWD 202  
Db 120 ISFORVLGICHPLASHKHGRGRRANVCGVWVLATACQLPTAVFAATGQIRNRVCYD 179  
QY 203 MTSQENFDRFAYGIVLTLSGFL-----SLIGHFGVLTQDQEPDQARGPEHEDRQHSPPSQ 258  
Db 180 LSPPIILSTRVLPYGMALTVIGFLFELPALLACVCMARRLCRODGPAGPVAQERRSKAAR 239  
QY 259 VPHDHTGVWPHLPFCALPYHSLLLPHLL-SAFSGLPALDGSQGLQDMEASGEQOL 317  
Db 240 M-----AVVVAFAISFLPHITKTAYLAVRSTPGVSCFVLETFFAAYKGR 287  
QY 318 PQPS-----PVLSE-----KGGKNRVRLLOKL 339  
288 PFASVNSVLDPLIFYFTQOKFRQPHDLQRL 319  
RESULT 3  
US-08-513-974B-56  
; Sequence 56, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236356  
; FILING DATE: 30-SEP-1994

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189274  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189273  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189272  
; FILING DATE: 11-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Resnick, David S.  
; REGISTRATION NUMBER: 34,235  
; REFERENCE/DOCKET NUMBER: 45753  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-513-974B-56  
Query Match 20.3%; Score 392.5; DB 3; Length 328;  
Best Local Similarity 33.0%; Pred. No. 3.9e-27;  
Matches 109; Conservative 42; Mismatches 142; Indels 37; Gaps 9;  
QY 33 DMNYSQEGGL-----COFSEYKQVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSGATT 88  
Db 4 DNGTQALGALGPTTCVYRENFKOLLPPVYSAVLAAGLPLNICVITQICTSRRLATTAV 63  
QY 89 YLVNLMVADLLYVL-LPFLIITVSLDORWPFGLLCKLVHFLVINYLGSTLLTCTSVH 147  
Db 64 YTLNLAADLLYACSLPLIYNYAQQDHWPFGLFACRLVRFYANLHGSILFLTCTISFQ 123  
QY 148 QFLGVCHPLCSLPYR-TRRHAMLGSTTWALVVLQLLPTLAFSHDYINGOMIWDMTSQ 206  
Db 124 RYLGHCHPLAPWHKGRGRRANVCGVWVLATVTTQCLPTAFAATGQIRNRVCYDLSPP 183  
QY 207 ENFDRFPAYGIVLTLSGFL-----SLIGHFGVLTQDQEPDQARGPEHEDRQHSPPSQVHPD 262  
Db 184 ALATHYMPYGMALTVIGFLFELPALLACVCLLAC-----RLCRQDQGAEPVAQ 231  
QY 263 HPTGVNPLHPLFCALPYHSLLLPHLL-SAFSGLPALDGSQGLQDMEASGECEQOLPQP- 320  
Db 232 ERGKAARMVAVVAAAFASISFLPHITKTAYLAVGSTPGVPCVTV--LEAFAAAYKGRPF 289  
QY 321 -----SPVLSF-----KGGKNRVRLLOKL 339  
Db 290 ASANSVLDPLIFYFTQOKFRRRPHELLQKL 319  
RESULT 4  
US-08-513-974B-380  
; Sequence 380, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA



APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1945  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 372:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-513-974B-372

Query Match 19.8%; Score 383.5; DB 3; Length 327;  
Best Local Similarity 40.9%; Pred. No. 2.4e-26;  
Matches 83; Conservative 32; Mismatches 81; Indels 7; Gaps 4;  
QY 29 MEKVDNMTSQEQL----COFSEKYKOVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWS 84  
DB 1 MEQ-DNGTQAPGLPTTCVYREDEKRLLLTPVSVVVGVLNLCVIAQICASRRTLT 59  
QY 85 CATTYLVNLMVADLLYL-LPFLIITYSLDDRPFGELCKLVHFLFYINLYSGILLTLCISVHOFGLGVCHPLCSLPY 143  
DB 60 RSAYVTNLALADLACSLPLLIYNYARCDHWPFGLACRFYRFLFYANLHGSILFTC 119  
QY 144 ISVHQFLGVCHPLCSLPYR-TRRHAWLGTSTTVALVQLLPTLAFSHTDYINGQMIWYD 202  
DB 120 ISFQVLGICHPLASHMKRGRRAANVCGVWVLATVTAQCLRTAVTAATGIQRNRCYD 179  
QY 203 MTSQENFDRLFAYGIVTLTGFL 225  
DB 180 LSPPIILSTRYPYGMALTIVIGFL 202

RESULT 6  
US-08-442-134A-2  
Sequence 2, Application US/08442134A  
Patent No. 5596088  
GENERAL INFORMATION:  
APPLICANT: Boucher, Richard C.  
APPLICANT: Weisman, Gary A.  
APPLICANT: Turner, John T.  
APPLICANT: Harden, Thomas K.  
APPLICANT: Parr, Claude E.  
APPLICANT: Sullivan, Daniel M.  
APPLICANT: Erb, Laura  
APPLICANT: Lustig, Kevin D.  
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and  
TITLE OF INVENTION: Null Cells Expressing P2U Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell, Seltzer, Park & Gibson  
STREET: Post Office Drawer 34009  
CITY: Charlotte  
STATE: No. 5596088th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,134A  
FILING DATE: 16-MAY-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sibley, Kenneth D.  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5470-71A  
TELEPHONE: 919-420-2200  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-442-134A-2  
Query Match 19.3%; Score 374.5; DB 1; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.8e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;  
QY 43 COFSEKYKOVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWCATTYLVNLMVADLLYL 102  
DB 25 CRNEDEKYVLLPVSYGVVGVLCGLNAVGLYIFLCRLKTNASTTYMFLHAYSDALYAA 84  
QY 103 -LPFLIITYSLDDRPFGELCKLVHFLFYINLYSGILLTLCISVHOFGLGVCHPLCSLPY 161  
DB 85 SLPLLVIYVYARGDHPFSTVCLCKLVRFLEYTNLYCSILFTLCISVHRCGLGVLRPLSLRW 144  
QY 162 RTRRHAWLGTSTTVALVQLLPTLAFSHTDYINGQMIWYDMSQENFDRLFAYGIVTL 221  
DB 145 GRARYARRVAGAVVWLACQAPVLYEVVTS-ARGPLTCHDTSAPELFSRFVAYSSYM-- 201  
QY 222 SGFLSLGLHGFVLF-----TDQEPDQARGEPHEDRQHSPOVHP 261  
DB 202 ---LGLL--FAVFAYLVYCVLMARLLKPAVGTSG-----GLPRAKKSVRT---- 245  
QY 262 DHTGVPWPLHPLF-CALPYH-----SLLPHHLLSAFSGLPALDGSQCGLODME 309  
DB 246 ----TAVLAVFALCFPLPHVTITLYYSFRSLDLSCHTLNAIN-----MAYKVTRLA 293  
QY 310 ASGECEQLPQSPVLSFKGKNRVLQKLRQNKLGHPA-GRKRCPLNRS 360  
DB 294 SANSC-----LDPVLYLAGQLRVRFARDAKP-PTGPSPATPARRTGLRRS 339  
RESULT 7  
US-08-444-581B-2  
Sequence 2, Application US/08444581B  
Patent No. 5607836  
GENERAL INFORMATION:  
APPLICANT: Boucher, Richard C.  
APPLICANT: Weisman, Gary A.  
APPLICANT: Turner, John T.  
APPLICANT: Harden, Thomas K.  
APPLICANT: Parr, Claude E.  
APPLICANT: Sullivan, Daniel M.  
APPLICANT: Erb, Laura  
APPLICANT: Lustig, Kevin D.  
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and  
TITLE OF INVENTION: Null Cells Expressing P2U Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell, Seltzer, Park & Gibson  
STREET: Post Office Drawer 34009  
CITY: Charlotte  
STATE: No. 5607836th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; ZIP: 28234
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC Compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/446,088A
;
; FILING DATE: 19-MAY-1995
;
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Kenneth D. Sibley
;
; REGISTRATION NUMBER: 31,665
;
; REFERENCE/DOCKET NUMBER: 5470-71C
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 919-420-2200
;
; TELEFAX: 919-881-3175
;
; INFORMATION FOR SEQ ID NO: 2:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 375 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; US-08-446-088A-2

```

[illegible]

```

RESULT 9
US-08-559-524A-3
; Sequence 3, Application US/08559524A
; Patent No. 5871963
; GENERAL INFORMATION:
; APPLICANT: Conley, Pamela B.
; APPLICANT: Jantzen, Hans-Michael
; TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/559,524A  
FILING DATE: 15-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044481-5010-00-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-559-524A-3

Query Match 19.3%; Score 374.5; DB 2; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.8e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;  
QY 43 CFSEKIKOVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYYL 102  
Db 25 CRNEDEKIVLLPVSYGVVCLGLCLNAVGLYFICRLKTNASTYMFHLAVSDALYAA 84  
QY 103 -LPFLIITYSLDRWPFGEELCKLVHFLYINLYSGILLTICISVHQFGLVCHPLCSLPY 161  
Db 85 SLPLLYVYARGDHPFSTVLCVLRFLYTNLYCSILFCLTICISVHRCGLVRLRLSLRW 144  
QY 162 RTTRHAWLGTSITWLVQLPLTAFSHTDYINGOMIWDTSQENFDRFLFAYGIVLTL 221  
Db 145 GRARYARRVAGVWLVLAQAPVLYFVTTS-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201  
QY 222 SGFLSLHGHFGLVLF-----TDGQEPDQARGPHEPDROHSPSOVHP 261  
Db 202 ---LGLL---FAVPFVILVCYVLMARRLLKPAYGTSG-----GLPRAKRSVRT---- 245  
QY 262 DHTPTGWPLHPLF--CALPYH-----SLLPHLLLSAFSGPLALDGSQGLQDME 309  
Db 246 -----IAVVLAVFALCFPLFHVTRTYLVSFRSLDLSCHTLNAIN-----NAYKVTRLA 293  
QY 310 ASGECEQLPQSPVLSFKGKNRVRLLQKLRONKLGHEPA-GRKRCPLNRS 360  
Db 294 SANSC-----LDPVLYFLAGQLRVRFARDAKP-PTGSPSPATPARRTLGLRRS 339

RESULT 10  
US-08-749-707-3  
Sequence 3, Application US/08749707  
Patent No. 6063582  
GENERAL INFORMATION:  
APPLICANT: Conley, Pamela B.  
APPLICANT: Jantzen, Hans-Michael  
TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/749,707  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044481-5010-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-749-707-3  
Query Match 19.3%; Score 374.5; DB 3; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.8e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;  
QY 43 CFSEKIKOVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYYL 102  
Db 25 CRNEDEKIVLLPVSYGVVCLGLCLNAVGLYFICRLKTNASTYMFHLAVSDALYAA 84  
QY 103 -LPFLIITYSLDRWPFGEELCKLVHFLYINLYSGILLTICISVHQFGLVCHPLCSLPY 161  
Db 85 SLPLLYVYARGDHPFSTVLCVLRFLYTNLYCSILFCLTICISVHRCGLVRLRLSLRW 144  
QY 162 RTTRHAWLGTSITWLVQLPLTAFSHTDYINGOMIWDTSQENFDRFLFAYGIVLTL 221  
Db 145 GRARYARRVAGVWLVLAQAPVLYFVTTS-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201  
QY 222 SGFLSLHGHFGLVLF-----TDGQEPDQARGPHEPDROHSPSOVHP 261  
Db 202 ---LGLL---FAVPFVILVCYVLMARRLLKPAYGTSG-----GLPRAKRSVRT---- 245  
QY 262 DHTPTGWPLHPLF--CALPYH-----SLLPHLLLSAFSGPLALDGSQGLQDME 309  
Db 246 -----IAVVLAVFALCFPLFHVTRTYLVSFRSLDLSCHTLNAIN-----MAYKVTRLA 293  
QY 310 ASGECEQLPQSPVLSFKGKNRVRLLQKLRONKLGHEPA-GRKRCPLNRS 360  
Db 294 SANSC-----LDPVLYFLAGQLRVRFARDAKP-PTGSPSPATPARRTLGLRRS 339

RESULT 11  
US-08-513-974B-373  
Sequence 373, Application US/08513974B  
Patent No. 6114139  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
APPLICANT: Hosoya, Masaki  
APPLICANT: Fujii, Ryo  
APPLICANT: Ohtaki, Tetsuya  
APPLICANT: Fukusumi, Shoji  
APPLICANT: Ohgi, Kazuhiro  
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
NUMBER OF SEQUENCES: 380  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/513.974B  
FILING DATE: 14-SEP-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01599  
FILING DATE: 10-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995  
APPLICATION DATA:  
APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 373:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-513-974B-373

Query Match 19.1%; Score 369; DB 3; Length 373;  
Best Local Similarity 31.6%; Pred. No. 5.4e-25;  
Matches 113; Conservative 42; Mismatches 115; Indels 88; Gaps 12;

Qy 43 CQSEKYKQVYLSAYSIIFILGLPLNGTVLWHSWGTWKWSCATTYLVNLMVADLLVYL 102  
Db 25 CRNEDEKQVLLPVSYGVCVGLCLNVALYIFLCRLKTNASTTTFWFLAVSDSLYAA 84  
Qy 103 -LPFLIITYSLDDRWPEGELLCKLVHFLFYINLYGSIILLTCTISVHQVLGCHPLCSLPY 161  
Db 85 SLPLLVYVYARGDHWPESTVLCKLVRELFVNTLYCSILFLTCTISVHRCGLVLRPLSLRW 144  
Qy 162 RTRRHWLGTSTTVALVQLLPLTAFSHDTYINGQMIWDTMTSQENFDRLFYAGIVLTL 221  
Db 145 GRARYARRVAVVWVLVLAQAPVLYFTTSVTRTCHDTSARELFSHFVAYSSVY-- 202  
Qy 222 SGFLSLGLHFGVLF-----TDQEPDQARGEPEHEDRQ--HSPSQV 259  
Db 203 --LGLL--FAVPSVILVCVLMARLLKPAYCTTG-----GLPRKRKSVRTIALV 250

Qy 260 HPDHPTGVWPLHPLFCALPYH-----SLLPLPHLLSAFS-----GLPALDGSQCQ 304  
Db 251 -----LAVFAL-----CFLPHVTRTLTYVSFRSLDLSCHTLNAINMAYKITRPLASANS- 300  
Qy 305 LODMEASGECEQLPQPSVLSFKGCKNRVRLLOKLR-----QNKLGEPHAGR 351  
Db 301 -----LDPVLYFLAGQRLVRFARDAKPPTEPTPSQARRKLGRLHPRNR 343

RESULT 12  
US-08-459-046-2  
Sequence 2, Application US/08459046  
Patent No. 6008039  
GENERAL INFORMATION:  
APPLICANT: Coleman, Roger  
APPLICANT: Au-Young, Janice  
APPLICANT: Stuart, Susan G. 6008039el Human Purinergic P2U Receptor  
TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3330 Hillview Avenue  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,046  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PF-0038 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 328 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-046-2

Query Match 19.0%; Score 368.5; DB 3; Length 328;  
Best Local Similarity 31.8%; Pred. No. 5.1e-25;  
Matches 105; Conservative 42; Mismatches 146; Indels 37; Gaps 9;

Qy 33 DMNFSQOGL-----CQSEKYKQVYLSAYSIIFILGLPLNGTVLWHSWGTWKWSCATT 88  
Db 4 DNGTDQALGPLPTTCVYRENFQKLLPPVYSVAVLAPALPLNICVITQITCSRRALTAV 63  
Qy 89 YLVNLMVADLLVYL-LPFLIITYSLDDRWPEGELLCKLVHFLFYINLYGSIILLTCTISVH 147  
Db 64 YTLMLALPDLIYACSLPLLIYNAQGDHWPFGDFACRLVRFYANLHGRILFLTCTISFQ 123  
Qy 148 QFLGVCHPLCSLPYR-TRRHAWLGTSTTVALVQLLPLTAFSHDTYINGQMIWDTMTSQ 206  
Db 124 RYLGLCHPLAPWHKRGGRRAWLVCVAVLAVTTQCLPTAFATGIGQRNRTVCYDLSPP 183  
Qy 207 ENFDRLFYAGIVLTLSCFL-----SLLGHFGLVFTDGOEPDQARGEPEHEDRQHSQVHPD 262  
Db 184 ALATHYMPYGMALTVIGFLPLFAALLACYCILLAC-----RLCRDQGAEPVAVQ 231  
Qy 263 HPTGVWPLHPLFCALPYHSLLLPHLL-SAFSGPLALDGSQCGLQDMEASGECEQLPQP- 320



```

; REFERENCE/DOCKET NUMBER: 044481-5010-00-US
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 202-467-7000
;

```

[illegible]

Search completed: May 22, 2003, 12:51:36  
Job time : 17 secs

```

RESULT 15
US-08-749-707-4
; Sequence 4, Application US/08749707
; Patent No. 6063582
; GENERAL INFORMATION:
; APPLICANT: Conley, Pamela B.
; APPLICANT: Jantzen, Hans-Michael
; TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/749,707
; APPLICATION NUMBER: US/08/749,707
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 044481-5010-01-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-749-707-4
Query Match 16.9% Score 326.5; DB 3; Length 373;

```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:49:50 ; Search time 22 Seconds  
(without alignments)  
1622.825 Million cell updates/sec

Title: US-10-023-586B-2  
Perfect score: 1936  
Sequence: 1 MLSTLLPSRSGSRGAL.....QNKLGHPAGKRCFCPLNRS 360

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1189	61.4	338	9	US-09-885-453-4
2	921	47.6	170	9	US-09-782-974C-14
3	371	19.2	341	9	US-10-270-587-3
4	370	19.1	97	9	US-09-791-932-92
5	367	19.0	365	9	US-09-077-173A-2
6	346	17.9	362	9	US-10-092-135-3
7	346	17.9	362	9	US-10-092-135-4
8	346	17.9	362	9	US-09-779-679-28
9	329.5	17.0	373	9	US-10-092-135-7
10	328	16.9	373	9	US-10-092-135-6
11	326.5	16.9	373	9	US-10-092-135-5
12	320	16.5	299	9	US-10-270-144-4
13	281.5	14.5	336	9	US-09-782-974C-86
14	278.5	14.4	337	9	US-10-023-775B-2
15	278.5	14.4	337	9	US-10-270-144-2
16	278.5	14.4	337	9	US-10-188-405-8
17	278.5	14.4	337	9	US-09-885-453-1
18	278.5	14.4	337	10	US-09-943-798-4
19	274.5	14.2	302	9	US-10-024-494-30

20	273.5	14.1	344	9	US-10-024-494-8
21	250.5	12.9	367	9	US-09-828-478-6
22	249	12.9	339	9	US-09-828-478-4
23	249	12.9	339	10	US-09-848-889-12
24	249	12.9	339	10	US-09-788-133-2
25	248.5	12.8	391	9	US-09-990-940-20
26	246	12.7	347	10	US-09-823-114-24
27	242.5	12.5	317	9	US-09-891-138A-2
28	242.5	12.5	359	9	US-10-094-417-10
29	242.5	12.5	359	9	US-09-782-974C-76
30	242.5	12.5	359	10	US-09-739-151-2
31	242	12.5	359	9	US-09-867-915-3
32	242	12.5	359	9	US-10-218-574-9
33	242	12.5	359	10	US-09-867-569-9
34	240.5	12.4	391	9	US-09-841-720-8
35	239.5	12.4	359	9	US-10-190-469-1
36	239.5	12.4	366	9	US-09-779-679-25
37	236	12.2	795	10	US-09-254-783A-1
38	236	12.2	795	12	US-10-152-058-1
39	234.5	12.1	276	10	US-09-943-798-2
40	234.5	12.1	362	9	US-09-850-948-2
41	234	12.1	359	10	US-09-966-871-81
42	234	12.1	359	12	US-10-039-645-81
43	233	12.0	427	9	US-10-007-132-4
44	233	12.0	427	9	US-09-771-287-4
45	232.5	12.0	334	10	US-09-765-034-2

ALIGNMENTS

RESULT 1

US-09-885-453-4  
; Sequence 4, Application US/09885453  
; Publication No. US20030088080A1  
; GENERAL INFORMATION:  
; APPLICANT: Communi, Didier  
; TITLE OF INVENTION: RECEPTOR GPCRx10  
; FILE REFERENCE: 9409/2082  
; CURRENT APPLICATION NUMBER: US/09/885,453  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/885,453  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Amino acid sequence  
; LOCATION: (1)..(338)  
; OTHER INFORMATION: GPCRx6 amino acid sequence  
US-09-885-453-4

Query Match	61.4%	Score 1189;	DB 9;	Length 338;
Best Local Similarity	98.7%;	Pred. No. 2.4e-106;		
Matches	225;	Conservative	0;	Mismatches 3;
Indels	0;	Gaps	0;	
QY	1	MLSTLLPSRSGSRGALLLEGASRDMEKVDNMTSQEQGLQFSEKVKQVYLSLAYS	60	
Db	1	MLSTLLPSRSGSRGALLLEGASRDMEKVDNMTSQEQGLQFSEKVKQVYLSLAYS	60	
QY	61	IFILGLPLNGTIVLWHSNGQTKRWSGATTVLVNLMVADLLVLLPFLITYSLDDRWPFGE	120	
Db	61	IFILGLPLNGTIVLWHSNGQTKRWSGATTVLVNLMVADLLVLLPFLITYSLDDRWPFGE	120	
QY	121	LLCKLVHFLFYINLYGSIILLITCISVHQFLGVCHPCLSLPYTRRHAWLGSTTVALVVL	180	
Db	121	LLCKLVHFLFYINLYGSIILLITCISVHQFLGVCHPCLSLPYTRRHAWLGSTTVALVVL	180	
QY	181	QLLEPTLAFSHTDYINGOMIYDMT SQENFDRLFAIGVILTLGFLSLL	228	
Db	181	QLLEPTLAFSHTDYINGOMIYDMT SQENFDRLFAIGVILTLGFLSLL	228	

Db 181 QLLPTLAFSHTDYINGOMIWDMTSQENFDRFLFAYGIVLTLSGFFPSL 228

RESULT 2

US-09-782-974C-14  
; Sequence 14, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor  
; FILE REFERENCE: 411USPHRM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/165,838  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 60/198,568  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-974C-14

Query Match 47.6%; Score 921; DB 9; Length 170;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 83 WSCATTYLVNLMVADLLYVLLPFLIITYSLDDRWPFGLCKLVHFLFYINLYGSILLLT 142  
Db 1 WSCATTYLVNLMVADLLYVLLPFLIITYSLDDRWPFGLCKLVHFLFYINLYGSILLLT 60  
143 CISVHQFLGVCHPLCSLYPRTRRHAWLTSTTVALVQLLPTLAFSHTDYINGOMIWD 202  
Db 61 CISVHQFLGVCHPLCSLYPRTRRHAWLTSTTVALVQLLPTLAFSHTDYINGOMIWD 120  
Qy 203 MTSQENFDRFLFAYGIVLTLSGFLSLGLHFGVLFTDQEPDQARGEPHEDR 252  
Db 121 MTSQENFDRFLFAYGIVLTLSGFLSLGLHFGVLFTDQEPDQARGEPHEDR 170

RESULT 3

US-10-270-587-3  
; Sequence 3, Application US/10270587  
; Publication No. US20030054487A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Yi  
; TITLE OF INVENTION: Human G-Protein Coupled Receptor  
; FILE REFERENCE: PF217C2  
; CURRENT APPLICATION NUMBER: US/10/270,587  
; CURRENT FILING DATE: 2002-10-16  
; PRIOR APPLICATION NUMBER: US 09/908,593

; PRIOR FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 08/781,456  
; PRIOR FILING DATE: 1997-01-10  
; PRIOR APPLICATION NUMBER: US 60/009,902  
; PRIOR FILING DATE: 1996-01-11  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-270-587-3

Query Match 19.2%; Score 371; DB 9; Length 341;  
Best Local Similarity 31.6%; Pred. No. 1.5e-27;  
Matches 113; Conservative 42; Mismatches 115; Indels 88; Gaps 12;  
Qy 43 CQSEKQVYVLSLAYSIIFILGLPLNGTVLWHSWGQTKRWCATTYLVNLMVADLLYVL 102  
Db 22 CRFNEDEKYLVPVSYGVVCLNVALYIFLCRLKTNASTYMFHLAVSDSLAA 31  
Qy 103 -LPFLIITYSLDDRWPFGLCKLVHFLFYINLYGSILLLTCTISVHQFLGVCHPLCSLY 161  
Db 82 SLPLLYVYARGDHWPFSTVLCKLVRFYINLYCSILFLCTISVHRCLGLVRLPLHSLRW 141  
Qy 162 RTRRHAWLTSTTVALVQLLPTLAFSHTDYINGOMIWDYDMT SQENFDRFLFAYGIVLT 221  
Db 142 GRARYARRAAVAVVVLACQAPVLYFVTTSVRGTRITCHDT SARELFSHFVAYSSVM-- 199  
Qy 222 SGFLSLGLHFGVLP-----TDGEPDQARGEPHEDRQ--HSPQOV 259  
Db 200 ---LGLL--FAVPFVILVCYVLMARLLKPAYGTG-----GLPRAKRKSVRTIALV 247  
Qy 260 HPDHTGVWPLHPLFCALPYH-----SLLPFHLLSAFS-----GLPALDGSOCG 304  
Db 248 -----LAVFAL----CFLPFHVTRTYYSFRSLDLSCHTLNAINMAYKITRPLASANS- 297  
Qy 305 LODMEASGECEQLPQSPVLSFKGKNRVLQLKLR-----ONKLGEPHAPGR 351  
Db 298 -----LDPVLYFLAGQLRVREARDAKPTEPTSPQARRKLGHRPNR 340

RESULT 4

US-09-791-932-92  
; Sequence 92, Application US/09791932  
; Publication No. US200300303451A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Parodi, Luis A.  
; APPLICANT: Hiebsch, Ronald R.  
; APPLICANT: Lind, Peter  
; APPLICANT: Kaytes, Paul S.  
; APPLICANT: Ruff, Valerie  
; APPLICANT: Huff, Rita M.  
; APPLICANT: Wood, Linda S.  
; TITLE OF INVENTION: No. US200300303451A1el G Protein-Coupled Receptors Cross-Ref  
; FILE REFERENCE: 00325.US1  
; CURRENT APPLICATION NUMBER: US/09/791,932  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/184,305  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,304  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,303  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,397  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,247  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/188,880  
; PRIOR FILING DATE: 2000-03-13  
; PRIOR APPLICATION NUMBER: 60/217,369  
; PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/217,370  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/218,492  
PRIOR FILING DATE: 2000-07-20  
PRIOR APPLICATION NUMBER: 60/186,810  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/188,064  
PRIOR FILING DATE: 2000-03-09  
PRIOR APPLICATION NUMBER: 60/186,457  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: 60/213,861  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: 60/194,344  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: 60/218,337  
PRIOR FILING DATE: 2000-07-14  
NUMBER OF SEQ ID NOS: 184  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 92  
LENGTH: 97  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-932-92

Query Match 19.1%; Score 370; DB 9; Length 97;  
Best Local Similarity 100.0%; Pred. No. 3.5e-28;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLIPLSRGSGRSGRGALLLEGASRDMKVKDMNTSQEGLCOFSEKQYVLSIYLSI 60  
DB 24 MSLIPLSRGSGRSGRGALLLEGASRDMKVKDMNTSQEGLCOFSEKQYVLSIYLSI 83  
QY 61 IFILGLPLNGTVLM 74  
DB 84 IFILGLPLNGTVLM 97

RESULT 5  
US-09-077-173A-2  
Sequence 2, Application US/09077173A  
Publication No. US20030082674A1  
GENERAL INFORMATION:  
APPLICANT: RECEPTOR AND NUCLEIC ACID MOLECULE ENCODING  
TITLE OF INVENTION: SAID RECEPTOR  
NUMBER OF SEQUENCES: 4  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0; Version #1.30 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,173A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/BE 96/00123  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-173A-2

Query Match 19.0%; Score 367; DB 9; Length 365;  
Best Local Similarity 41.1%; Pred. No. 3.9e-27;  
Matches 79; Conservative 30; Mismatches 75; Indels 8; Gaps 3;

QY 43 CFSEKQYVLSIYLSIIFILGLPLNGTVLWHSWGOTKRWSCATTYLVNLMVADLLYL 102  
DB 27 CWFDEPKFILLPVSAVVFVLGGLNAPTLWLFIFRLRPMDATATYMFHLASDITLVL 86.

QY 103 -LPFLIITYSLDDRWPFGEGLLCKLVHFLFYINLYGSIILLTCTISVHOFGLGVCHPLCLSPY 161  
DB 87 SLPTLIIVYAAHNHWPGEIGKVFVFLFYVNLVCSVLELTCTISVHRYLGICHPRLALRW 146  
QY 162 RTRHAWLGTSTTVALVVLQLLPTLAFSHDYDINGQMIWDTMTSQENFDRFLFAYGIVLTL 221  
DB 147 GRPLAGLLCLAVNLVYVAGCLVPNLFVTTSNKGTTVLCHDTTRPEEFDHVHFS----- 201  
QY 222 SGFLSLILGHFGV 233  
DB 202 SAVMGLL--FGV 211

RESULT 6  
US-10-092-135-3  
Sequence 3, Application US/10092135  
Publication No. US20030054374A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
FILE REFERENCE: D0134.NP  
CURRENT APPLICATION NUMBER: US/10/092,135  
CURRENT FILING DATE: 2002-03-06  
PRIOR APPLICATION NUMBER: US 60/273,808  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/278,983  
PRIOR FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 3  
LENGTH: 362  
TYPE: PRT  
ORGANISM: GALLUS GALLUS  
US-10-092-135-3

Query Match 17.9%; Score 346; DB 9; Length 362;  
Best Local Similarity 30.5%; Pred. No. 4.1e-25;  
Matches 97; Conservative 49; Mismatches 128; Indels 44; Gaps 10;

QY 53 YLSLAYSIIIFILGLPLNGTVLWHSWGOTKRWSCATTYLVNLMVADLLYL-LPFLIITYS 111  
DB 42 YLPTVYILVITGFLGNSVAIMFVFMHMPWSGISVYMFNLALADFLYLTLPALIFYFF 101  
QY 112 LDDRWPFGEGLLCKLVHFLFYINLYGSIILLTCTISVHOFGLGVCHPLCLSPYRTRHAWLGT 171  
DB 102 NKTDWIFGDVWCKLQRFIFHVNLYGSIILFCTISVHRYTGWHVHPLKSLGRLLKKNAYVS 161  
QY 172 STTVALVVLQLLPTLAFSHDYDINGQMI-WYDMTSQENFDRFLFAYGIVLTL-----SGFLS 226  
DB 162 SLVWALVAVIAPIILFYSGTGVRNKTITCYDTTADDEVLSYFVYSCTTVFMFCIPFIV 221  
QY 227 LLGHFGVLFTDGOEPDQARGEPHEDRHSQVHIPDHTGVWPHLPFLCALPYHSL-LLP 285  
DB 222 ILGCYGLI-----VKALIYKDLNLSPLRRK-----SYLVIIVITVFAVSYP 264  
QY 286 HHLLSAFSGPLALD---GSOCGLQD-----MEASGECEQLPQSPVLSFKGKN 331  
DB 265 FHVMTLNLRARLDFQPMCAFNDKVYATYQVTRGLASLNSC-----VDPILYLAGDT 319  
QY 332 -RVRLQLKLRQNLGEHP 348  
DB 320 FRRLSRATRKSSRRSEP 337

RESULT 7  
US-10-092-135-4  
Sequence 4, Application US/10092135  
Publication No. US20030054374A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
FILE REFERENCE: HGPRBMV27

```
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092.135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 362
; TYPE: PRT
; ORGANISM: MELEAGRIS GALLOPAVO
; US-10-092-135-4

Query Match      17.9%  Score 346; DB 9; Length 362;
Best Local Similarity 30.5%; Pred. No. 4.1e-25;
Matches 97; Conservative 49; Mismatches 128; Indels 44; Gaps 10;

Qy 53 YLSLAYSIIIFILGPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYL-LPFLIIITYS 111
Db 42 YLPTVYILVFTIGTGLNSVAIWMFVHMRPWSGISVYMFNLALADFLVLTLPALIFYF 101
Qy 112 LDDRPFCEELCKLVHFLFYINLYGSILLTLCISVHQFLGCHPLCSLPYTRRHAWLGT 171
Db 102 NKTWIFGDVNMCKLQRFIFHVNLYGSILFLTCISVHRYTGVVHPLKSLGRKKNAVYS 161
Qy 172 STTWALVYQLLPTLAFSHTDYINGQMT-WYDMTSQENFDRFLFAYGIVLT- 226
Db 162 SLVWALVAVAPIILFYSGTGVRNKTITCYDTTADFLRSFYVSMCTTFVMECIPIV 221
Qy 227 LLGHFGVLFTDQBPQARGPEHEDRQHSQVHPDHTGVWPLHPLFCALPYHSL-LLP 285
Db 222 ILGCYGLI-----VKALIYKDLNLSPLRRK-----SIYLVIIIVLTVEAVSYLP 264
Qy 286 HLLLSAFSGLPALD---GSCQGLD-----MEASGECEQLPQSPVLSFKGKN 331
Db 265 FHVWKTLLRLARLDFQTPQMCAPNDKVYATQVTRGLASLNSC-----VDPIYFLAGDT 319
Qy 332 -RVRLQLKRONKLGHP 348
Db 320 FRRRLSRATRKSSRRSEP 337

RESULT 8
US-09-779-679-28
; Sequence 28, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1e1 Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USN 60/219758
; PRIOR FILING DATE: 2000-07-20
```

```
; PRIOR APPLICATION NUMBER: USN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-09-779-679-28

Query Match      17.9%  Score 346; DB 9; Length 362;
Best Local Similarity 30.5%; Pred. No. 4.1e-25;
Matches 97; Conservative 49; Mismatches 128; Indels 44; Gaps 10;

Qy 53 YLSLAYSIIIFILGPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYL-LPFLIIITYS 111
Db 42 YLPTVYILVFTIGTGLNSVAIWMFVHMRPWSGISVYMFNLALADFLVLTLPALIFYF 101
Qy 112 LDDRPFCEELCKLVHFLFYINLYGSILLTLCISVHQFLGCHPLCSLPYTRRHAWLGT 171
Db 102 NKTWIFGDVNMCKLQRFIFHVNLYGSILFLTCISVHRYTGVVHPLKSLGRKKNAVYS 161
Qy 172 STTWALVYQLLPTLAFSHTDYINGQMT-WYDMTSQENFDRFLFAYGIVLT- 226
Db 162 SLVWALVAVAPIILFYSGTGVRNKTITCYDTTADFLRSFYVSMCTTFVMECIPIV 221
Qy 227 LLGHFGVLFTDQBPQARGPEHEDRQHSQVHPDHTGVWPLHPLFCALPYHSL-LLP 285
Db 222 ILGCYGLI-----VKALIYKDLNLSPLRRK-----SIYLVIIIVLTVEAVSYLP 264
Qy 286 HLLLSAFSGLPALD---GSCQGLD-----MEASGECEQLPQSPVLSFKGKN 331
Db 265 FHVWKTLLRLARLDFQTPQMCAPNDKVYATQVTRGLASLNSC-----VDPIYFLAGDT 319
Qy 332 -RVRLQLKRONKLGHP 348
Db 320 FRRRLSRATRKSSRRSEP 337

RESULT 9
US-10-092-135-7
; Sequence 7, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPT
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092.135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 373
; TYPE: PRT
; ORGANISM: RATTUS NORVEGICUS
; US-10-092-135-7

Query Match      17.0%  Score 329.5; DB 9; Length 373;
```

Best Local Similarity 36.7%; Pred. No. 1.7e-23;  
Matches 77; Conservative 36; Mismatches 82; Indels 15; Gaps 4;  
QY 53 YLSLAYSIIIFILGLPLNGTVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 53 YLPNAVILYFIIGFLGNSVAIMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDRWPFGELLCKLVHFLFYINLYGSILLTICISVHQFLGVCHPLGSLPYRTTTHAWLGT 171  
Db 113 NKTDWIFGDVMCKLQRFIFHVNLYGSILFETCSAHRYSGVVPYPLKSLGRLKKNNAIYVS 172  
QY 172 STTWALVVLQLLPTLAFSHTDY-INQMTIWDYMTSOENFDRFYAGVITLSGF-----LS 226  
Db 173 VLWLVVVAISPILFYSGTGIRKKNKTVTCYDSTDSYLSRYFIYSMCTTVMFCVPLVL 232  
227 LLGHFGVLTGDEPQAGEPHEDRQHS 256  
Db 233 ILGCYGLI-----VRALIYKLDLNSP 253

RESULT 10  
US-10-092-135-6  
; Sequence 6, Application US/10092135  
; Publication No. US20030054374A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
; FILE REFERENCE: D0134.NP  
; CURRENT APPLICATION NUMBER: US/10/092,135  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: US 60/273,808  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 60/278,983  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 6  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-10-092-135-6

Query Match 16.9%; Score 326; DB 9; Length 373;  
Best Local Similarity 29.9%; Pred. No. 2.3e-23;  
Matches 93; Conservative 51; Mismatches 123; Indels 44; Gaps 10;  
QY 53 YLSLAYSIIIFILGLPLNGTVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 53 YLPNAVILYFIIGFLGNSVAIMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDRWPFGELLCKLVHFLFYINLYGSILLTICISVHQFLGVCHPLGSLPYRTTTHAWLGT 171  
Db 113 NKTDWIFGDVMCKLQRFIFHVNLYGSILFETCSAHRYSGVVPYPLKSLGRLKKNNAIYVS 172  
QY 172 STTWALVVLQLLPTLAFSHTDYINGOMI-WYDWTSOENFDRFYAGVITLSGF-----LS 226  
Db 173 VLWLVVVAISPILFYSGTGIRKKNKTVTCYDSTDSYLSRYFIYSMCTTVMFCVPLVL 232  
QY 227 LLGHFGVLTGDEPQAGEPHEDRQHS 256  
Db 233 ILGCYGLI-----VRALIYKLDLNSP 253  
QY 286 HLLSAFSGPLALD---GSCGGLQD-----MEASGECEQLPQSPVLSFKGKN 331  
Db 276 FHVAKTNLRLARLDFTQTPANCAFNDRVYATYQTRGLASLNSC-----VDPILYFLAGDT 330  
QY 332 -RVRLLOKLRQ 341  
Db 331 FRRRLSRATRK 341  
RESULT 11

US-10-092-135-5  
; Sequence 5, Application US/10092135  
; Publication No. US20030054374A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
; FILE REFERENCE: D0134.NP  
; CURRENT APPLICATION NUMBER: US/10/092,135  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: US 60/273,808  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 60/278,983  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 5  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: BOS TAURUS  
US-10-092-135-5

Query Match 16.9%; Score 326.5; DB 9; Length 373;  
Best Local Similarity 36.7%; Pred. No. 3.2e-23;  
Matches 77; Conservative 35; Mismatches 83; Indels 15; Gaps 4;  
QY 53 YLSLAYSIIIFILGLPLNGTVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 53 YLPNAVILYFIIGFLGNSVAIMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDRWPFGELLCKLVHFLFYINLYGSILLTICISVHQFLGVCHPLGSLPYRTTTHAWLGT 171  
Db 113 NKTDWIFGDVMCKLQRFIFHVNLYGSILFETCSAHRYSGVVPYPLKSLGRLKKNNAIYVS 172  
QY 172 STTWALVVLQLLPTLAFSHTDYINGOMI-WYDWTSOENFDRFYAGVITLSGF-----LS 226  
Db 173 VLWLVVVAISPILFYSGTGIRKKNKTVTCYDSTDSYLSRYFIYSMCTTVMFCVPLVL 232  
QY 227 LLGHFGVLTGDEPQAGEPHEDRQHS 256  
Db 233 ILGCYGLI-----VRALIYKLDLNSP 253

RESULT 12  
US-10-270-144-4  
; Sequence 4, Application US/10270144  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO00750CON  
; CURRENT APPLICATION NUMBER: US/10/270,144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 299  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-270-144-4

Query Match 16.5%; Score 320; DB 9; Length 299;  
Best Local Similarity 38.3%; Pred. No. 1e-22;  
Matches 72; Conservative 33; Mismatches 77; Indels 6; Gaps 3;  
QY 53 YLSLAYSIIIFILGLPLNGTVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 12 YLPNAVILYFIIGFLGNSVAIMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 71

```
Qy 112 LDDRWPFGEELCKLVHFLFYINLYSGSILLTLCISVHOFGLVCHPLCSLPYTRRRHAWGT 171
Db 72 NKTDFGDMCKLQRFTHFNLYSGSILFLTCISAHRYSGVVYPLKSLGRLKKNAIYVS 131
Qy 172 STTWALVVLQPLTLAFSHT-DYINGOMIWDMTSQENFDRFLFAYGIVLTLSGF-----LS 226
Db 132 VLVWLVVVAISPILFYSGTGTRKNKTVCTDYTTSNDLYRSFYISMCTTVAMFCIPLVL 191
Qy 227 LLGHFGLV 234
Db 192 ILGCYGLI 199

RESULT 13
US-09-782-974C-86
; Sequence 86, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 411USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
9-782-974C-86

Query Match 14.5%; Score 281.5; DB 9; Length 336;
Best Local Similarity 33.3%; Pred. No. 6.1e-19;
Matches 59; Conservative 37; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSTIFTLGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKSSIIIMNLACTDLLYLTSLPFLIH 91
Qy 109 TYSLDDRWPFGEELCKLVHFLFYINLYSGSILLTLCISVHOFGLVCHPLCSLPYTRRRHAW 168
Db 92 YYASGENWIFGDFMCKIRSFHFNLYSSILFLTCFSIFRYCVIIHPMSCFSIHKTRCAV 151
Qy 169 LGTSTTWALVVLQPLP-TLAFSHTDYINGOMIWDMTSQENFDRFLFAYGIVLTLSGF 224
Db 152 VACAVVMIISLVAVIPMTFLTITSTNRTN-RSACLDTSSDELNTIKWYNLILTASTF 207

RESULT 14
US-09-782-974C-86
; Sequence 86, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 411USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
9-782-974C-86

Query Match 14.5%; Score 281.5; DB 9; Length 336;
Best Local Similarity 33.3%; Pred. No. 6.1e-19;
Matches 59; Conservative 37; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSTIFTLGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKSSIIIMNLACTDLLYLTSLPFLIH 91
Qy 109 TYSLDDRWPFGEELCKLVHFLFYINLYSGSILLTLCISVHOFGLVCHPLCSLPYTRRRHAW 168
Db 92 YYASGENWIFGDFMCKIRSFHFNLYSSILFLTCFSIFRYCVIIHPMSCFSIHKTRCAV 151
Qy 169 LGTSTTWALVVLQPLP-TLAFSHTDYINGOMIWDMTSQENFDRFLFAYGIVLTLSGF 224
Db 152 VACAVVMIISLVAVIPMTFLTITSTNRTN-RSACLDTSSDELNTIKWYNLILTASTF 207

RESULT 14
US-10-023-775B-2
; Sequence 2, Application US/10023775B
; Publication No. US20030022282A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (EP(GB) only)
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)
; APPLICANT: Fidoock, Mark David
; TITLE OF INVENTION: No. US20030022282A1el Polypeptide
; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-775B-2

Query Match 14.4%; Score 278.5; DB 9; Length 337;
Best Local Similarity 32.8%; Pred. No. 1.2e-18;
Matches 58; Conservative 38; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSTIFTLGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKSSIIIMNLACTDLLYLTSLPFLIH 91
Qy 109 TYSLDDRWPFGEELCKLVHFLFYINLYSGSILLTLCISVHOFGLVCHPLCSLPYTRRRHAW 168
Db 92 YYASGENWIFGDFMCKIRSFHFNLYSSILFLTCFSIFRYCVIIHPMSCFSIHKTRCAV 151
Qy 169 LGTSTTWALVVLQPLP-TLAFSHTDYINGOMIWDMTSQENFDRFLFAYGIVLTLSGF 224
Db 152 VACAVVMIISLVAVIPMTFLTITSTNRTN-RSACLDTSSDELNTIKWYNLILTATTF 207

RESULT 15
US-10-270-144-2
; Sequence 2, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
US-10-270-144-2

Query Match 14.4%; Score 278.5; DB 9; Length 337;
Best Local Similarity 32.8%; Pred. No. 1.2e-18;
Matches 58; Conservative 38; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSTIFTLGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKSSIIIMNLACTDLLYLTSLPFLIH 91
```









[illegible]



[illegible]

N;Alternate names: G-protein coupled receptor  
C;Species: Homo sapiens (man)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999  
C;Accession: T09508  
R;Bohm, S.K.; Trumpff, A.; Khtin, L.M.; Kong, W.; Payan, D.G.; Bunnett, N.W.  
submitted to the EMBL Data Library, April 1997  
A;Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retinol  
A;Reference number: Z16705  
A;Accession: T09508  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-344 <BOH>  
A;Cross-references: EMBL:AF000546; NID:g2232068; PID:g2232069  
C;Genetics:  
A;Map position: 13  
Superfamily: ATP receptor P2u  
Keywords: G protein-coupled receptor; transmembrane protein

Query Match 14.2%; Score 274.5; DB 2; Length 344;  
Best Local Similarity 33.8%; Pred. No. 1.1e-16;  
Matches 70; Conservative 39; Mismatches 77; Indels 21; Gaps 7;  
7;

QY 32 VDMNTSQEQGLQFSEKQVLSLAYSIIFILGPLNGTVLWHSWGQTKRWSCATTYLV 91  
Db 2 VSNSSH-----CFYNSFKYTLGCMFVMVLGVSNCVAIYIFICVLKVRNETTYMI 57  
QY 92 NLMVADLLVYL-LPFLITYSLDDRWPFGLLCKLVHFLVINYLGSIILLTCTISVHOFL 150  
Db 58 NLAUSDLLFVETLPRFP-FYTRRNWFFGLLCKISVWLYTNYGSIILFELTCTISVDRFL 116  
QY 151 GVCHPLCSLPYTRRRHAWLTSTTWALVLOLPT--LAFSHTDYINGOMIWDYDMSQEN 208  
Db 117 AIVYPFKSKTLTRKNAKIVCTGVMLTVIGGSAPAVEVQSTHSGNNASEACE-----EN 171  
QY 209 FD-----RLFAYGIVLTLSGLSLLGHF 231  
Db 172 FPEATWYTLRSRVI-----FIEIVGFF 194

RESULT 11  
JC5549  
heptahelical P2Y5-like receptor - human  
C;Species: Homo sapiens (man)  
C;Date: 02-Sep-1997 #sequence\_revision 05-Sep-1997 #text\_change 24-Sep-1999  
C;Accession: JC5549  
R;Janssens, R.; Boeynaens, J.M.; Godart, M.; Communi, D.  
Biochem. Biophys. Res. Commun. 236, 106-112, 1997  
A;Title: Cloning of a human heptahelical receptor closely related to the P2Y5 receptor.  
A;Reference number: JC5549; MUID:97366605; PMID:9223435  
A;Accession: JC5549  
A;Molecule type: DNA  
A;Residues: 1-370 <JAN>  
A;Cross-references: DDBJ:AF005419; NID:g2240034; PIDN:AAB66372.1; PID:g2240035  
C;Superfamily: ATP receptor P2u

Query Match 13.4%; Score 258.5; DB 2; Length 370;  
Best Local Similarity 33.2%; Pred. No. 3.1e-15;  
Matches 69; Conservative 35; Mismatches 93; Indels 11; Gaps 4;

QY 28 DMKVDNMTSQEGL-----CFSEKQVLSLAYSIIFILGPLNGTVLWHSNGQ 79  
Db 8 DFQFQDSNSSLRPLRGNATANTTCIVDDSEKYNLNGAVISVVFILGNTNSVLSLFVFCFR 67  
QY 80 TKRWSCATTYLVNLMVADLLVYL-LPFLITYSLDDRWPFGLLCKLVHFLVINYLGSI 138  
Db 68 MKMRSETAFTNLAVSDLLFVCTLPKIP-FYFNRRHWPFGDTLCKISGTAFLNITGSM 126  
QY 139 LLLFCTISVHOFLGVCHPLCSLPYTRRRHAWLTSTTWALVLOLPTLAFSHTDYINGOM 198  
Db 127 LFLFCTISVDRFLAIVYFPRSTIRTRNSAIVCAGVILVLSGGISASLFTTWNVNTAT 186  
QY 199 IWDYDMSQENFDR-LFAYGIVLTLSGFL 225

Db 187 TCPEGLSKRVWKTYLSKITIFIEVVGEI 214

#### RESULT 12

JC2492

G protein-coupled receptor 1 - rat

N;Alternate names: GPR-1

C;Species: Rattus norvegicus (Norway rat)

C;Date: 16-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 08-Oct-1999

C;Accession: JC2492

R;Marchese, A.; Cheng, R.; Lee, M.C.; Porter, C.A.; Heiber, M.; Goodman, M.; George,

Biochem. Biophys. Res. Commun. 205, 1952-1958, 1994

A;Title: Mapping studies of two G protein-coupled receptor genes: An amino acid diffe

A;Reference number: JC2492; MUID:95110347; PMID:7811287

A;Accession: JC2492

A;Molecule type: mRNA

A;Residues: 1-353 <MAR>

A;Cross-references: GB:S74702; NID:g786483; PIDN:AAB32978.1; PID:g786484

C;Superfamily: vertebrate rhodopsin

C;Keywords: glycoprotein; lipoprotein; phosphoprotein; receptor; thiolester bond; tra

F;74-94/Domain: transmembrane #status predicted <TM2>

F;112-133/Domain: transmembrane #status predicted <TM3>

F;134-135/Region: DR motif

F;155-175/Domain: transmembrane #status predicted <TM4>

F;209-229/Domain: transmembrane #status predicted <TM5>

F;246-266/Domain: transmembrane #status predicted <TM6>

F;295-306/Domain: transmembrane #status predicted <TM7>

F;14-273/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;150,231/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status pred

F;330/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 13.3%; Score 258; DB 2; Length 353;

Best Local Similarity 30.8%; Pred. No. 3.2e-15;

Matches 72; Conservative 41; Mismatches 81; Indels 40; Gaps 8;

QY 26 SRDM--EKVDNMTSQEQGLQFSEKQVLS-----LAYSIFILGPLNGTVLW-- 74

Db 4 SREMLFEELDNYSYALEYYSQEPDAENYVPGIVHVSLLLYALAFVLGIPGNAIVWF 63

QY 75 -HSMGQTKRWSCATTYLVNLMVADLLVYL-LPFLITYSLDDRWPFGLLCKLVHFLFYI 132

Db 64 GFKWKKT----VTTLWFNLNLAIDFIFVFLPLYSIVALSFFHWPFGWLCKLSFIAQL 119

QY 133 NLXGSILLTCTISVHOFLGVCHPLCSLPYTRRRHAWLTSTTWALVLOLPTLAFSHTD 192

Db 120 NMESVFFELTVISLDRVILHILHPLGSLPHRTLKNLSLLVLFVWLLASLLGGPTLYFRDTV 179

QY 193 YINGQMTWYDMSQENFDRL-----FAYGIVLT-----SGFLSLL 228

Db 180 EVNNRIICYN-----NFQVEYELTMRHHLVLTWVAKFLGYLLPLLTMTSSCYLC LI 228

#### RESULT 13

A41795

somatostatin receptor 1 - human

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 24-Nov-1999

C;Accession: A41795

R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.

Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992

A;Title: Cloning and functional characterization of a family of human and mouse somat

A;Reference number: A41795; MUID:92108031; PMID:1346068

A;Accession: A41795

A;Molecule type: DNA

A;Residues: 1-391 <YAM>

A;Cross-references: GB:M81829; NID:g307433; PIDN:AAA58247.1; PID:g307434

A;Note: sequence extracted from NCBI backbone (NCBIN:74767, NCBIIP:74768)

C;Genetics:

A;Gene: GDB:SSTR1

A;Cross-references: GDB:134185; OMIM:182451

A;Map position: 14q13-14q13

A;Introns: #status absent

C;Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; phd  
F:58-84/Domain: transmembrane #status predicted <TM1>  
F:95-120/Domain: transmembrane #status predicted <TM2>  
F:132-153/Domain: transmembrane #status predicted <TM3>  
F:173-195/Domain: transmembrane #status predicted <TM4>  
F:220-230/Domain: transmembrane #status predicted <TM5>  
F:269-296/Domain: transmembrane #status predicted <TM6>  
F:302-326/Domain: transmembrane #status predicted <TM7>  
F:4,44,48,381/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:130-208/Disulfide bonds: #status predicted  
F:172/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predicted  
F:265/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predicted  
F:339/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 12.8%; Score 248.5; DB 2; Length 391;  
Best Local Similarity 32.6%; Pred. No. 2.5e-14;  
Matches 74; Conservative 41; Mismatches 103; Indels 9; Gaps 6;

Qy 7 PSRGS--RSGSRGALLLEGASRDMEKVDMMNTSQOGLCFSEKQYKQVLSIYFIIL 64  
Db 16 PSPGCGEGGSGRGP--GAGAADGMEEPGRNASQNTLSE--GQGSAILISFIYSVVCILV 71

Qy 65 GLPLNGTVLHWSWGOTKRWSCATTVLVNLMVAD-LLYVLLPFLIITYSLDDRWPFCELLC 123  
Db 72 GLCGNSMVIYVILRYAKMKTATNIYILNLAIDEILMLSVPEL-VTSTLLRHWPFGALLC 130

Qy 124 KLVHFLFVINYLSILLTLCISVHQFLGVCHPLCPSLPTRRHAWLGSTTWALVVLQLL 183  
Db 131 RLVLSDVADVNMTSIYCLTVLSVDYVAVVHPKARVRRPTVAKVNVGVVLSLLVIL 190

Qy 184 PTLAFSHDTYINGQMIWDMTSQENFDR-LFAYGTVLTLSGFLSLG 229  
Db 191 PTWFSRTAANSDDGTAVACNMLMPEAQRWLVGFVLYTFLMGFLPVG 237

RESULT 14  
S15403  
angiotensin II receptor type 1 - bovine  
C:Species: Bos primigenius indicus x Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Nov-1999  
C:Accession: S15403  
R:Sasaki, K.; Yamano, Y.; Bardhan, S.; Iwai, N.; Murray, J.J.; Hasegawa, M.; Matsuda, Y.  
Nature 351, 230-233, 1991  
A:Title: Cloning and expression of a complementary DNA encoding a bovine adrenal angiotensin II receptor type 1  
A:Reference number: S15403; MUID:91251900; PMID:2041569  
A:Accession: S15403  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-359 <SAS>  
A:Cross-references: GB:X62294; NID:g43; PIDN:CAA44182.1; PID:g44  
C:Superfamily: vertebrate rhodopsin

Query Match 12.8%; Score 247.5; DB 2; Length 359;  
Best Local Similarity 31.0%; Pred. No. 2.8e-14;  
Matches 63; Conservative 42; Mismatches 87; Indels 11; Gaps 4;

Qy 34 MNTSQOGL-----CFSEKQYKQVLSL--AYSIIIFILGLPLNGTVLHWSWGOTKRWSC 85  
Db 3 LNSSTDGDKIRTDGCPKAGRHNYIFIMPTLYSIIFVVGFGNSLVVIVFYMKLKV 62

Qy 86 ATTYLVNLMVADLLYL-LPFLIITYSLDDRWPFCELLCKLVHFLFYINLYGSILLTLCI 144  
Db 63 ASVFLNLALADCLFLLPLPLVAVVAVTAMEYRWPFGNYLCKIASASVSFNLYASVELTCL 122

Qy 145 SVHQFLGVCHPLCPSLPTRRHAWLGSTTWALVVLQLLPTLAFSHDTYINGQMIWDMT 204  
Db 123 SIDRYLAIVHPKSLRRTMLVAKVTCIIIIWLLAGLASLPTIIHNRVFFIENTNITVCAF 182

Qy 205 SQENFDRLFAYGIVLT--LSGFL 225  
Db 183 HYESQNSTLPVGLGTTKNLGL 205

RESULT 15  
I48705  
proteinase activated receptor 2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Nov-1999  
C:Accession: I48705  
R:Nystedt, S.; Larsson, A.K.; Aberg, H.; Sundelin, J.  
J. Biol. Chem. 270, 5950-5955, 1995  
A:Title: The mouse proteinase-activated receptor-2 cDNA and gene. Molecular cloning a  
A:Reference number: I48705; MUID:95197620; PMID:7890726  
A:Accession: I48705  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-399 <RES>  
A:Cross-references: EMBL:Z48043; NID:9663020; PIDN:CAA88097.1; PID:9663021  
C:Superfamily: ATP receptor P2u

Query Match 12.7%; Score 246; DB 2; Length 399;  
Best Local Similarity 35.1%; Pred. No. 4.3e-14;  
Matches 68; Conservative 34; Mismatches 80; Indels 12; Gaps 7;

Qy 48 KYQVYLSIAYSIIFILGLPLNGTVLHWSWGOTKRWSCATTVLVNLMVADLLYVLLPFLI 107  
Db 74 KLTTFVLPVVIIVVIGLPSNGMALWFLFRTKKHFAVITMANLADLLSVIWFPLK 133

Qy 108 ITYSL-DDRWPFCELLCKLVHFLFYINLYGSILLTLCISVHQFLGVCHPLCPSLPTRRH 166  
Db 134 ISYHLHGNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPM-GHP-RKKN 191

Qy 167 AWLQTS-TTWALVVLQLLPTLAFSHDTYINGQMI--WYDMTSQENF-DRLFAYGIVLTL 222  
Db 192 IAVGVSLAIWLLIFLVITPLIYVMKOTIYPALNITTCHDVLPVEVLVGDMMENYFLSLAIG 251

Qy 223 GFLSLLGHGFLFT 236  
Db 252 VFL-----FPALLT 260

Search completed: May 22, 2003, 12:51:13  
Job time : 23 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:47:10 ; Search time 14 Seconds  
(without alignments)  
1066.534 Million cell updates/sec

Title: US-10-023-586B-2  
Perfect score: 1936  
Sequence: 1 MSLILPSRSGSRGAL.....QNKLGHPAGRKRCPLNRS 360

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	447	23.1	328	1 P2Y3 CHICK	Q98907 gallus gall
2	447	23.1	328	1 P2Y3 MELGA	Q93361 melagris g
3	416	21.5	537	1 P2Y8 XENLA	P79928 xenopus lae
4	400	20.7	328	1 P2Y6 RAT	O63371 rattus norv
5	388	20.0	361	1 P2Y4 RAT	O35811 rattus norv
6	387.5	20.0	328	1 P2Y6 HUMAN	Q15077 homo sapien
7	387.5	20.0	377	1 P2Y2 HUMAN	P41231 homo sapien
8	386	19.9	361	1 P2Y4 MOUSE	Q9JJ57 mus musculu
9	371	19.2	373	1 P2Y2 MOUSE	P35383 mus musculu
10	367	19.0	365	1 P2Y4 HUMAN	P51582 homo sapien
11	366.5	18.9	374	1 P2Y2 RAT	P41232 rattus norv
12	346	17.9	362	1 P2YR CHICK	P34996 gallus gall
13	346	17.9	362	1 P2YR MELGA	P49652 melagris g
14	329.5	17.0	373	1 P2YR RAT	P49651 rattus norv
15	328	16.9	373	1 P2YR HUMAN	P47900 homo sapien
16	326.5	16.9	373	1 P2YR BOVIN	P48042 bos taurus
17	320	16.5	373	1 P2YR MOUSE	P49650 mus musculu
18	282.5	14.6	308	1 P2Y5 CHICK	P32250 gallus gall
19	274.5	14.2	344	1 P2Y5 HUMAN	P43657 homo sapien
20	260.5	13.5	165	1 P2Y4 CRIGR	P58826 cricetus
21	258.5	13.4	370	1 P2Y9 HUMAN	Q99677 homo sapien
22	258	13.3	353	1 GPRI RAT	P46090 rattus norv
23	252	13.0	353	1 APJ XENLA	P79960 xenopus lae
24	250.5	12.9	367	1 GP17 HUMAN	Q13304 homo sapien
25	250.5	12.9	374	1 PAR3 HUMAN	O00254 homo sapien
26	249	12.9	397	1 PAR2 RAT	Q63645 rattus norv
27	248.5	12.8	359	1 AG2R SHEEP	O77590 ovis aries
28	248.5	12.8	391	1 SSRI HUMAN	P30872 homo sapien
29	247.5	12.8	359	1 AG2R BOVIN	P25104 bos taurus
30	246	12.7	399	1 PAR2 MOUSE	P55086 mus musculu
31	245.5	12.7	391	1 SSRI RAT	P28646 rattus norv
32	244.5	12.6	359	1 AG2R CANFA	P43240 canis famil
33	244.5	12.6	391	1 SSRI MOUSE	P30873 mus musculu

RESULT 1				
P2Y3_CHICK	34	242.5	12.5	397
ID	P2Y3_CHICK	STANDARD;	PRT;	328 AA.
AC	Q98907;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).			
GN	P2Y3.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96319774; PubMed=8700132;			
RA	Webb T.E., Henderson D., King B.F., Wang S., Simon J.,			
RA	Bateson A.N., Burnstock G., Barnard E.A.;			
RT	"A novel G protein-coupled P2 purinoceptor (P2Y3) activated			
RT	preferentially by nucleoside diphosphates.";			
RL	Mol. Pharmacol. 50:258-265(1996).			
CC	-1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADP > UTP > ATP = UDP. THE			
CC	ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE			
CC	A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X98283; CAA66930.1; -.			
DR	HSSP; P34996; 1DDO.			
DR	InterPro: IPR000276; GPCR_Rhodpsn.			
DR	Pfam: PF00001; 7tm1; 1.			
DR	PRINTS; PR00237; GPCRHHOOPS.			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.			
DR	PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein.			
FT	DOMAIN 1 23			
FT	TRANSMEM 23 43			
FT	DOMAIN 44 57			
FT	TRANSMEM 58 78			
FT	DOMAIN 79 96			
FT	TRANSMEM 97 117			
FT	DOMAIN 118 139			
FT	TRANSMEM 140 160			
FT	DOMAIN 161 189			
FT	TRANSMEM 190 210			
FT	DOMAIN 211 231			

P55085 homo sapien  
P30556 homo sapien  
P47749 xenopus lae  
P30555 sus scrofa  
Q13725 homo sapien  
P34976 oryctolagus  
P29089 rattus norv  
P29754 mus musculu  
P25085 rattus norv  
Q92081 rattus norv  
Q9W08 mus musculu  
Q96991 homo sapien

ALIGNMENTS

```

FT TRANSMEM 232 252 6 (POTENTIAL).
FT DOMAIN 253 275 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 276 298 7 (POTENTIAL).
FT DOMAIN 299 323 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 94 172 BY SIMILARITY.
SQ SEQUENCE 328 AA; 37586 MW; 7A3BFIC91F54FAAB CRC64;

Query Match 23.1%; Score 447; DB 1; Length 328;
Best Local Similarity 45.9%; Pred. No. 2.1e-22;
Matches 85; Conservative 31; Mismatches 67; Indels 2; Gaps 2;

QY 43 CQFSKYQVLSLAYSIIFILGLPLNGTVLHWSGQTKRWSCATTYLVNLMVLLVYL 102
DB 13 CTFHEEFQVLLPLVYSVVFLLGLPLNAVIGQIWLARKALTRTTIYMLNADLLVYC 72
QY 103 -LPFLIITYSLDDRPFGELCKLVHFLFYINLYGSLILLCISVHQFLGVCPLCSL-P 160
DB 73 SLPLLIYNTQKDYWPFQDFTCKFRFYTNLHGSILFLTCISVQRYMGICHLASWHK 132
QY 161 YTRRHAWLGTSTTVALVQLLPTLAFSHTDYINGOMIYDMTSQENFDRLFAFGVILT 220
DB 133 KKGKLLTWLCAAVFVIAQCLPTFVFASTGTQRNRTVCYDLSPPDRSTSYFPYGITLT 192
221 LSGFL 225
193 IINGFL 197

RESULT 2
P2Y3_MEIGA
ID P2Y3_MEIGA STANDARD; PRT; 328 AA.
AC O93361;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).
GN P2Y3.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98401046; PubMed=9730913;
RT Li Q., Olesky M., Palmer R.K., Harden T.K., Nicholas R.A.;
RT "Evidence that the p2y3 receptor is the avian homologue of the
RL mammalian p2y6 receptor.";
RL Mol. Pharmacol. 54:541-546(1998).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UDP< ADP = UTP. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF069555; AAC23863.1; -
CC HSSP; P34996; 1DDD.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1.1.
CC PRINTS; PR00237; GPCRRHODPSN.
CC PROSITE; PS00237; G-PROTEIN_RECP_F1_1; FALSE_NEG.
CC PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
CC DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 23 43 1 (POTENTIAL).

```

```

FT DOMAIN 44 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 2 (POTENTIAL).
FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 97 117 3 (POTENTIAL).
FT DOMAIN 118 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 4 (POTENTIAL).
FT DOMAIN 161 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 5 (POTENTIAL).
FT DOMAIN 211 231 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 232 252 6 (POTENTIAL).
FT DOMAIN 253 275 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 276 298 7 (POTENTIAL).
FT DOMAIN 299 323 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 94 172 BY SIMILARITY.
SQ SEQUENCE 328 AA; 37594 MW; B74D49B99C7164A5 CRC64;

Query Match 23.1%; Score 447; DB 1; Length 328;
Best Local Similarity 44.6%; Pred. No. 2.1e-22;
Matches 86; Conservative 32; Mismatches 73; Indels 2; Gaps 2;

QY 35 NTSBOGLCQFSEKYQVLSLAYSIIFILGLPLNGTVLHWSGQTKRWSCATTYLVNLM 94
DB 5 NTAGRNCTQEEFKQVLLPLVYSVVFLLGLPLNAVIGQIWLARKALTRTTIYMLNLA 54
QY 95 VADLLYL-LPFLIITYSLDDRPFGELCKLVHFLFYINLYGSLILLCISVHQFLGVC 153
DB 65 TADLLYVCSLPLLIYNTQKDYWPFQDFTCKFRFYTNLHGSILFLTCISVQRYMGIC 124
QY 154 HPLCSL-PYTRRHAWLGTSTTVALVQLLPTLAFSHTDYINGOMIYDMTSQENFDRL 212
DB 125 HPLASWHKKGKLLTWLCAAVFVIAQCLPTFVFASTGTQRNRTVCYDLSPPDRSASY 184
QY 213 FAGVILTLGFL 225
DB 185 FPGITLTITGFL 197

RESULT 3
P2Y8_XENLA
ID P2Y8_XENLA STANDARD; PRT; 537 AA.
AC P79928;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
GN P2Y purinoceptor 8 (P2Y8).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97284734; PubMed=9139711;
RT Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;
RT "Early expression of a novel nucleotide receptor in the neural plate
RL of Xenopus embryos.";
RL J. Biol. Chem. 272:12583-12590(1997).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, GTP AND ITP.
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

or send an email to licensee@lsb-sib.ch).
-----
CC EMBL; X99953; CAA68213.1; -.
CC DR HSP; P34996; 1DDD.
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam; PF00001; 7tm_1; 1.
CC DR PRINTS; PR00237; GPCRHOPOPSN.
CC DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
CC DR PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 50 70 1 (POTENTIAL).
CC FT DOMAIN 71 79 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 80 100 2 (POTENTIAL).
CC FT DOMAIN 101 118 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 119 139 3 (POTENTIAL).
CC FT DOMAIN 140 161 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 162 182 4 (POTENTIAL).
CC FT DOMAIN 183 210 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 211 231 5 (POTENTIAL).
CC FT DOMAIN 232 254 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 255 275 6 (POTENTIAL).
CC FT DOMAIN 276 292 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 293 316 7 (POTENTIAL).
CC FT DOMAIN 317 337 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 116 193 BY SIMILARITY.
CC FT CARBOHYD 26 26 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 29 29 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 537 AA; 62024 MW; B2CF24812F3C19F2 CRC64;

Query Match 21.5%; Score 416; DB 1; Length 537;
Best Local Similarity 31.8%; Pred. No. 3.3e-20;
Matches 110; Conservative 59; Mismatches 139; Indels 38; Gaps 10;

QY 31 KVDNN-TSQGGLCFSEKVKQVLSLAYSIIFILGLPLNGTVLWHSWGOTKRWCATTY 89
DB 22 KLLMNLNTEDEICVDFEGFKLLPVSYSAVFMVGLPLNIAAWIFIAKMRPNPTTVY 81

QY 90 LVNLMVADLLYVL-LPELIITYSLDDRWPEGLCKLVHFLYINLYGSIILLTLCISVHQ 148
DB 82 MFNLALSDTLVLSLPTLVVYYADKNWPEGLVKLVRELFYANLYSSILFTLCISVHR 141

QY 149 FLGVCHPLCSLPYTRRHANLGTSTWALVQLPLTAFSHDYINGQMIWDMTSQEN 208
DB 142 YRGVCHPITSLRRNNAKHAYVICALVWLSVTCLVPLNLFVTSVPKYNKTCIDTTRPED 201

QY 209 FDRLEFAY--GIVLTSGF--LSLLGHFGVLTGQEPDQARGEPHEDRHSQVHPDHP 264
DB 202 FARVEYSTAIMCLLFGIPCLTIAGCYGLMTRMLMKPIVS-----GNOQTLPYKKRSIK 256

QY 265 TGVNPLHPLFCALPYHSLLLPHHLLSAFSGLPALDGSQC-----GLQDMEASGE 313
DB 257 T-----IIFVMAFAICFEPFHITRTLYYYARLLGKICVLANVINYYKVTREPLASANS 310

QY 314 CEQLPQSPVLSF-KGKNRVRLLQKLURQKNLGEHPAGRKRCPLGN 358
DB 311 C-----IDPILYFLANDRYRRRLIRTVRRRS-----SVENRRRCMHTN 347

RESULT 4
P2Y6_RAT
ID P2Y6_RAT STANDARD; PRT; 328 AA.
AC Q63371;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE P2Y purinocceptor 6 (P2Y6).
GN P2Y6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

SEQUENCE FROM N.A.
TISSUE=Aortic smooth muscle;
MEDLINE=96064682; PubMed=7592819;
Chang K., Hanaoka K., Kumada M., Takawa Y.;
'Molecular cloning and functional analysis of a novel P2 nucleotide
receptor.';
J. Biol. Chem. 270:26152-26158(1995).
-!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UTP > ADP = 2-METHYLTHIO-ATP
> ADP-BETA-S > ATP = ATP-GAMMA-S. THE ACTIVITY OF THIS RECEPTOR IS
MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-
CALCIUM SECOND MESSENGER SYSTEM. FUNCTIONALLY COUPLED TO
PHOSPHOLIPASE C.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN VARIOUS TISSUES
INCLUDING LUNG, STOMACH, INTESTINE, SPLEEN, MESENTERY, HEART, AND,
MOST PROMINENTLY, AORTA.
-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensee@lsb-sib.ch).
-----
EMBL; D63665; BAA09816.1; -.
DR HSP; P34996; 1DDD.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 28 48 1 (POTENTIAL).
FT DOMAIN 49 62 2 (POTENTIAL).
FT TRANSMEM 63 83 3 (POTENTIAL).
FT DOMAIN 84 101 4 (POTENTIAL).
FT TRANSMEM 102 122 5 (POTENTIAL).
FT DOMAIN 123 144 6 (POTENTIAL).
FT TRANSMEM 145 165 7 (POTENTIAL).
FT DOMAIN 166 194 8 (POTENTIAL).
FT TRANSMEM 195 215 9 (POTENTIAL).
FT DOMAIN 216 236 10 (POTENTIAL).
FT TRANSMEM 237 257 11 (POTENTIAL).
FT DOMAIN 258 280 12 (POTENTIAL).
FT TRANSMEM 281 303 13 (POTENTIAL).
FT DOMAIN 304 328 14 (POTENTIAL).
FT DISULFID 99 177 BY SIMILARITY.
FT CARBOHYD 5 5 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 328 AA; 36677 MW; 58EBE1997038BA5B CRC64;

Query Match 20.7%; Score 400; DB 1; Length 328;
Best Local Similarity 32.5%; Pred. No. 2.3e-19;
Matches 108; Conservative 53; Mismatches 137; Indels 34; Gaps 9;

QY 29 MEKVDNMTSQEGL---COFSEKVKQVLSLAYSIIFILGLPLNGTVLWHSWGOTKRWS 84
DB 1 MER-DNGTIOAGLPPTTCVIREDFKRLLPVTVVGLPLNVCVIAQICASRRFLT 59

QY 85 CATTYLVNLMVADLLYVL-LPFLIITYSLDDRWPEGLCKLVHFLYINLYGSIILLTLC 143
DB 60 RSVAVYTLNLAADLLYACSLPLLIYNYARGDHPFGDLACFLVFLFVYANLHGSILFTC 119

QY 144 ISVHQFLGVCHPLCSLPYR-TRRHANLGTSTWALVQLPLTAFSHDYINGQMIWYD 202
DB 120 ISFQRLGICHPLAPWHKRGRRRAAVVCGVWLVVTAQCLPTAFVAAFGIQRNRTCYD 179

QY 203 MTSQENEDRLFAVGIIVLTLSGFL---SLIGHFGVLTGQEPDQARGEPHEDRHSQ 258
DB 180 LSPPIILSTRYLPYGMALTIVIGFLPFTALLACACYCRMARRLCRQDGPAGVPAQERRSAA 239

```

QY 259 VHPDPTGWPVLPFCALPYHSLLLPHLL-SAFSGLPALDGSQGLQDMASGECESOL 317  
 Db 240 M-----AVVAAVFISELPFHITKTAYLAVRSTPGVSCPVLTFFAAYKGR 287  
 QY 318 POPS-----PVLSF-----KGGKNRVRLLOKL 339  
 Db 288 PFASANSVLPILFYFTQOKFRQRPHDLLQKL 319

RESULT 5  
 P2Y4\_RAT  
 ID P2Y4\_RAT STANDARD; PRT; 361 AA.  
 AC O35811;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE P2Y purinoceptor 4 (P2Y4).  
 GN P2Y4 OR P2Y4  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP STRAIN=Sprague-Dawley; TISSUE=Liver;  
 Bogdanov F.D., Wildman S., King B.F., Burntack G.;  
 Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=98421785; PubMed=9751165;  
 RA Webb T.E., Henderson D., Roberts J.A., Barnard E.A.;  
 RT "Molecular cloning and characterization of the rat P2Y4 receptor.";  
 RL J. Neurochem. 71:1424-1434(1998).  
 CC 1- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that  
 activate a phosphatidylinositol-calcium second messenger system.  
 CC Not activated by ADP or UDP.  
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC 1- TISSUE SPECIFICITY: Widely expressed at low levels. In brain,  
 CC higher expression in the pineal gland and ventricular system.  
 CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y14705; CAA75007.1; -;  
 DR EMBL; Y11433; CAA72241.1; -;  
 DR HSSP; P34996; 1DD0.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm1.1;  
 DR PRINTS; PS00237; GPCRRHODPSN  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_F1\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN\_RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 30  
 FT TRANSMEM 31 58  
 FT DOMAIN 59 68  
 FT TRANSMEM 69 91  
 FT DOMAIN 92 108  
 FT TRANSMEM 109 127  
 FT DOMAIN 128 149  
 FT TRANSMEM 150 170  
 FT DOMAIN 171 192  
 FT TRANSMEM 193 218  
 FT DOMAIN 219 242  
 FT TRANSMEM 243 265  
 FT DOMAIN 266 283  
 FT TRANSMEM 284 305

FT DOMAIN 306 361 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 104 181 BY SIMILARITY.  
 FT CARBOHYD 175 175 N-LINKED (GLCNAC..)(POTENTIAL).  
 SQ SEQUENCE 361 AA; 40893 MW; 0377F96E54B449A3 CRC64;

Query Match 20.0%; Score 388; DB 1; Length 361;  
 Best Local Similarity 32.1%; Pred. No. 1.5e-18;  
 Matches 105; Conservative 54; Mismatches 126; Indels 42; Gaps 9;

QY 37 SOEGLCOFSEKYKOVYLSLAYSIIFILGLPLNGVTLVHWSGQTKRWSCATTLYNLMAVA 96  
 Db 17 SSGDGDRENEEFKILLPMSTAVVVLGLALNAPTLMLFLRLRPWDATATYMFHLALS 76  
 QY 97 DLLYVL-LPFLIITYSLDDRMPPFGELCKLVHFLFYINLYSGISILLTCTISVHQFGVCHP 155  
 Db 77 DTLYVLSLPTLVVYVAARNHWPFGTGLCKEVRFLFYWNLYCSVLFTCTISVHRYLGICH 136  
 QY 156 LCSLPYRTRRHAWLGTSTWALVQLLPTLAFSTHDYINGOMIWDYMTSQENFDR--LF 213  
 Db 137 LRATNGRPRFASLLCLGLVWVAGLVNLFVFTTNANGTTLCHDFTLPEFDHYVVF 196  
 QY 214 AYGIVLTLSG--FLSLGLHFGVLFDTGQEPDQARGEPHEDRQHSQVHPDHTGVPMLH 271  
 Db 197 SSANVLLFGLPFLITLVYCYGLMARRLYRPLPGAQ-----SSRLSLRTIAVLT 249  
 QY 272 PLFCALPYH---SLLPHHLLSAFSGLPALDGSQGL-----ODMEASGECESOL 317  
 Db 250 FAVCFVPFHITRTIYYQARLLQA-----DCHVLNVNVVYKVRPLASANS--- 296

QY 318 PQSPVL-SFKGKNRVRLLQKLQKRONK 343  
 Db 297 --LDPVLYFTGDKYRNQLQQLCRGSK 321

RESULT 6  
 P2Y6\_HUMAN  
 ID P2Y6\_HUMAN STANDARD; PRT; 328 AA.  
 AC Q15077; Q15754;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE P2Y purinoceptor 6 (P2Y6).  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Communi D., Parmentier M., Boeynaems J.M.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97432828; PubMed=9286708;  
 RA Somers G.R., Hammet F., Woollatt E., Richards R.I., Southey M.C.,  
 RA Venter D.J.;  
 RT "Chromosomal localization of the human P2y6 purinoceptor gene and  
 RT phylogenetic analysis of the P2y purinoceptor family.";  
 RL Genomics 44:127-130(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=97415792; PubMed=9268704;  
 RA Maier R., Glatz A., Mosbacher J., Bilbe G.;  
 RT "Cloning of P2Y6 cDNAs and identification of a pseudogene: comparison  
 RT of P2Y receptor subtype expression in bone and brain tissues.";  
 RL Biochem. Biophys. Res. Commun. 237:297-302(1997).  
 RN [4]  
 RP ERRATUM.  
 RX MEDLINE=98069816; PubMed=9412455;  
 RA Maier R., Glatz A., Mosbacher J., Bilbe G.;  
 RL Biochem. Biophys. Res. Commun. 240:298-302(1997).  
 CC 1- FUNCTION: RECEPTOR FOR EXTRACELLULAR UDP > UTP > ATP. THE ACTIVITY

OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.

-!- SUBCELLULAR LOCATION: Integral membrane protein.

-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X97058; CAA65770.1; --  
DR EMBL; U52464; AAB03572.1; --  
DR EMBL; AF007892; AAB80713.1; --  
DR EMBL; AF007891; AAB80712.1; --  
DR HSSP; P34996; 1DDD.  
DR Genew; HGNC:8543; P2RY6.  
DR MIM; 602451; --  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRRHODPSN.  
DR PROSITE; PS00237; G-PROTEIN\_RECP\_F1\_1; FALSE\_NEG.  
DR PROSITE; PS50262; G-PROTEIN\_RECP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 28 48 1 (POTENTIAL).  
FT DOMAIN 49 62 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 63 83 2 (POTENTIAL).  
FT DOMAIN 84 101 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 102 122 3 (POTENTIAL).  
FT DOMAIN 123 144 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 145 165 4 (POTENTIAL).  
FT DOMAIN 166 194 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 195 215 5 (POTENTIAL).  
FT DOMAIN 216 236 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 237 257 6 (POTENTIAL).  
FT DOMAIN 258 280 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 281 303 7 (POTENTIAL).  
FT DOMAIN 304 328 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 5 5 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT DISULFID 99 177 BY SIMILARITY.  
FT CONFLICT 2 3 MISSING (IN REF. 2).  
FT SEQUENCE 328 AA; 36429 MW; AAD6C55A43818107 CRC64;

Query Match 20.0%; Score 387.5; DB 1; Length 328;  
Best Local Similarity 32.7%; Pred. No. 1.5e-18;  
Matches 108; Conservative 42; Mismatches 143; Indels 37; Gaps 9;

QY 33 DMNTSQEGL----CFSEKVKQVYLSLAYSTIFILGLPLNGTIVLHWSWGOTKRWSCATT 88  
Db 4 DNGTQALGLPTTCVYRENFQALLPPVYSAVLAAGLPINLCVITQICTSRALTRAV 63  
QY YLVNLMVADLLVYL-LPELIITYSLDRWPFGEELCKLVFLFYINLYGSSILLITFCISVH 147  
Db 64 YTLNLALADLLVYACSLPLLYVYAQGDHPGDFACRLVFLFYANLHGSILFLTCISFQ 123  
QY 148 QFLGVCHPLCSLPYR-TRHAWLGTSTTVALVVLQLLPTLAFSHDYINGQMIWDMTSQ 206  
Db 124 RYLGIHPLAPWKKRGGRAWLVCVAVMLAVTTQCLPTAFAATGIQRNRTVCVYDLSP 183  
QY 207 ENFDRLFAYGIVLTLSGFL----SLIGHFGLVLTDSQEPDQARGEPHEDRHSQVHPD 262  
Db 184 ALATHYMPYGMALTVIGLLFPALALACVLLAC-----RLCRDGPAPV 231  
QY 263 HPTGVWPLHPLFCALPYHSLLPHPHLL-SAFSGPLPDGSCGLDMEASGEOLPOP- 320  
Db 232 ERGRKAARMVAVARAFSLFPFHTKTAYLAVRSTGVPCV--LEAFAAAYKGRPF 289  
QY 321 -----SPVLSF-----KGGKNRVRLQKL 339  
Db 290 ASANSVLPILFYFTQKFRFRPHELLQKL 319

RESULT 7

P2Y2\_HUMAN

ID P2Y2\_HUMAN STANDARD; PRT; 377 AA.

AC P41231; O96EM8;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)  
DE (Purinergic receptor).  
GN P2RY2 OR P2RU1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Airway epithelium;  
RX MEDLINE=94211846; PubMed=8159738;  
RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H., Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;  
RT "Cloning and expression of a human P2U nucleotide receptor, a target for cystic fibrosis pharmacotherapy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:3275-3279(1994).  
RN [2]  
RP REVISTONS  
RX MEDLINE=95108098; PubMed=7809171;  
RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H., Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;  
RT "Cloning and expression of a human P2U nucleotide receptor, a target for cystic fibrosis pharmacotherapy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:13067-13067(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Kidney, and Leukocyte;  
RA Strausberg, R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBSJ databases.  
CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that activate a phosphatidylinositol-calcium second messenger system. The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP = ADP.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND BRAIN.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; U07225; AAC04923.1; --  
DR EMBL; BC012104; AAH2104.1; --  
DR EMBL; BC028135; AAH28135.1; --  
DR PIR; A54946; A54946.  
DR HSSP; P34996; 1DDD.  
DR Genew; HGNC:8541; P2RY2.  
DR MIM; 600041; --  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRRHODPSN.  
DR PROSITE; PS00237; G-PROTEIN\_RECP\_F1\_1; 1.  
DR PROSITE; PS50262; G-PROTEIN\_RECP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 33 59 1 (POTENTIAL).  
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 71 93 2 (POTENTIAL).  
FT DOMAIN 94 110 EXTRACELLULAR (POTENTIAL).







```

QY 305 LODMEASGECEQLPQSPVLSFKGKNRRVLLQKLR-----QNKLGHPAGR 351
Db 301 -----LDPVLFLAGQLRVFARDAKPPTPTSPQARRKGLGHRNR 343

RESULT 10
P2Y4_HUMAN
ID P2Y4_HUMAN STANDARD; PRT; 365 AA.
AC P51582;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 4 (P2Y4) (Uridine nucleotide receptor) (UNR) (P2P).
GN P2RY4 OR NRU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96125055; PubMed=8537336;
RA Communi D., Piroton S., Parmentier M., Boeynaems J.-M.;
RT "Cloning and functional expression of a human uridine nucleotide
receptor.";
J. Biol. Chem. 270:30849-30852(1995).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96125054; PubMed=8537335;
RA Nguyen T., Erb L., Weisman G.A., Marchese A., Heng H.H.Q.,
RA Garrard R.C., George S.R., Turner J.T., O'Dowd B.F.;
RT "Cloning, expression, and chromosomal localization of the human
uridine nucleotide receptor gene.";
J. Biol. Chem. 270:30845-30848(1995).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96125053; PubMed=8537334;
RA Stam N.J., Klomp J., van der Heuvel M., Olijve W.;
RT "Molecular cloning and characterization of a novel orphan receptor
(P2P) expressed in human pancreas that shows high structural homology
to the P2U purinoceptor.";
FEBS Lett. 384:260-264(1996).
CC -!- FUNCTION: Receptor for UDP and UDP coupled to G-proteins that
activate a phosphatidylinositol-calcium second messenger system.
Not activated by ATP or ADP.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Pancreas.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL; X91852; CAA62963.1; -
DR EMBL; U40223; AAC50347.1; -
DR EMBL; X96597; CAA65415.1; -
DR HSSP; P34996; IDDD.
DR Gene; HGNC:8542; P2RY4.
DR MIM; 300038; -
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1.1; 1
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN_RECEPTOR_F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEPTOR_F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Polymorphism.
FT DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).
FT STRAIN 35 61 1 (POTENTIAL).
FT STRAIN 62 72 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 73 95 2 (POTENTIAL).
FT DOMAIN 96 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 131 3 (POTENTIAL).
FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 174 4 (POTENTIAL).
FT DOMAIN 175 196 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 197 222 5 (POTENTIAL).
FT DOMAIN 223 246 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 247 269 6 (POTENTIAL).
FT DOMAIN 270 287 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 288 309 7 (POTENTIAL).
FT DOMAIN 310 365 CYTOPLASMIC (POTENTIAL).
FT DISULFID 108 185 BY SIMILARITY.
FT VARIANT 168 168 V -> M (IN DBSNP:1152186).
FT VARIANT 178 178 N -> T (IN DBSNP:1152187).
FT VARIANT 191 191 P -> L (IN DBSNP:1152188).
FT CONFLICT 86 86 L -> V (IN REF. 2).
FT CONFLICT 234 234 S -> A (IN REF. 2).
SQ SEQUENCE 365 AA; 40963 MW; 23E0AFED3B7BDEED CRC64;

Query Match 19.0%; Score 367; DB 1; Length 365;
Best Local Similarity 41.1%; Pred. No. 3.4e-17;
Matches 79; Conservative 30; Mismatches 75; Indels 8; Gaps 3;

QY 43 COFSEKYOVLSLAYSIIFILGLPLNGVTLVHWSGOTKRWSCATYLVNLMVADLLYL 102
Db 27 CWFDEDFKILLPVSYAVVFLGLGNAPTLWLFTRLRPWDATATYMFHLASDLYVL 36
QY 103 -LPFLIITYSLDDRMPPFGELCKLVHFLFYINLYGSIILLTCTISVHQFLGVCHPLCSLPY 161
Db 87 SLPTLIYYVAAHNHPFEGTEICKFVRFYWNLYCSVFLTCTISVHYRLGICHPRLALRW 146
QY 162 RTRRHAWLGTSTTVALVQLLPTLAFSTDYINGQNIWDMTSQENFRLFAYGIVL 221
Db 147 GRPRLAGLLCLAVLWVAGCLVPLNFVTTNKGTTVLCHDTRPEEFDHVHFS----- 201
QY 222 SGFLSLLLGHFGV 233
Db 202 SAVMGLL--FGV 211

RESULT 11
P2Y2_RAT
ID P2Y2_RAT STANDARD; PRT; 374 AA.
AC P41232;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)
DE (Purnergic receptor).
GN P2RY2 OR P2RU1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95110548; PubMed=7811468;
RA Rice W.R., Burton F.M., Fiedel D.T.;
RT "Cloning and expression of the alveolar type II cell P2u-purnergic
receptor.";
J. Biol. Chem. 270:32732-32738(1995).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN-Wistar; TISSUE=pituitary;
RA Chen Z.P., Krull N., Xu S., Levy A., Lightman S.L.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX STRAIN-Wistar;

```



```

203  ---LGLL--FAVPFSIILVCYVLMARRLLKPAYGTTGLPRAKRKSVRTIALVJLVAFAL-- 255
273  LFCALPYH-----SLLPHHLLSAFS-----GLPALDGSQGLQDMASGGEQL 317
      |||:|      |||:|      |||:|      |||:|
256  --CFLPFHVRTLYYSFSLDLSCHTLNAINMAYKITRPLASANS----- 299
      |||:|      |||:|      |||:|      |||:|
318  PQSPVLSTFKGKNRVLLQKLR-----QNKLGHEHPAGR 351
      |||:|      |||:|      |||:|      |||:|
300  --LDPVLYFLAGORLVRFARDARKPATEPTPSQARRKGLGLHRPNR 342
      |||:|      |||:|      |||:|      |||:|

RESULT 12
P2YR_CHICK
ID P2YR_CHICK STANDARD; PRT; 362 AA.
AC P34996;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN P2RX1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
TI TISSUE=Brain;
RC MEDLINE=93285340; PubMed=8508924;
RA Webb T.E., Simon J., Krishek B.J., Bateson A.N., Smart T.G.,
RA King B.F., Burnstock G., Barnard E.A.;
RT "Cloning and functional expression of a brain G-protein-coupled ATP
RT receptor.";
RL FEBS Lett. 324:219-225(1993).
RN [2]
3D-STRUCTURE MODELING.
RP MEDLINE=97026278; PubMed=8872457;
RX van Rhee A.M., Fischer B., van Galen P.J.M., Jacobson K.A.;
RT "Modelling the P2Y purinoceptor using rhodopsin as template.";
RA Drug Des. Discov. 13:133-140(1995).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
CC INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, GASTROINTESTINAL TRACT,
CC SPLEEN AND LEG MUSCLE. IS NOT DETECTED IN THE HEART, LIVER,
CC STOMACH, LUNG AND KIDNEY.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; X73268; CAA51716.1; -.
DR PIR; S33733; S33733.
DR PDB; 1DDO; 11-JUL-96.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN
DR PROSITE; PS00337; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; 3D-structure.
FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 42 63 1 (POTENTIAL).
FT FT 64 76 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 77 98 2 (POTENTIAL).
FT DOMAIN 99 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 136 3 (POTENTIAL).

```

```

FT DOMAIN 137 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 177 4 (POTENTIAL).
FT DOMAIN 178 207 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 208 227 5 (POTENTIAL).
FT DOMAIN 228 254 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 255 274 6 (POTENTIAL).
FT DOMAIN 275 292 7 (POTENTIAL).
FT TRANSMEM 293 317 7 (POTENTIAL).
FT DOMAIN 318 362 CYTOPLASMIC (POTENTIAL).
FT DISULFID 113 191 BY SIMILARITY.
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 362 AA; 41194 MW; A806C88FB9514761 CRC64;

Query Match 17.9%; Score 346; DB 1; Length 362;
Best Local Similarity 30.5%; Pred. No. 7.7e-16;
Matches 97; Conservative 49; Mismatches 128; Indels 44; Gaps 10;

QY 53 YLSLAYSIFILGLPLNGTVLWHSWGQTKRWSGQKRWSCATYLVNLMVADLLVYL-LPFLIITYS 111
DB 42 YLPTVIIVITGFLGNSVAIMWFHMRPWSGISVYMFNLADFLVLTLPALIFYF 101

112 LDRWPFGECLCKLVHFLFYINLYGSILLTCTISVHOFGLGVCHPLCSLPYRTRRHAWLGT 171
DB 102 NKTDFWGDVCKLQRFIFHVNLYGSILLTCTISVHRVTGVVHPLKSLGRKKNAVYVS 161

QY 172 STTWALVQLLPTLAFSHDYNQMI-WYDWTSOENFDRFLFAYGIVLTL-----SGFLS 226
DB 162 SLVWALVAVIAPILFYSGTVGVRNKTTCYDTTADYLRSYFVSMCTTFVFCIPFIV 221

QY 227 LLGHGVLFTDQGEPDQARGEPEHEDRQSPSOVHPDPTGVVPLHPLFCALPYHSL-LLP 285
DB 222 ILGCVGLI-----VKALIYKDLNSPLRRK-----SIVLIIVLVFAVSYLP 264

QY 286 HLLSAFSGLPALD---GSQCGLQD-----MEASGECEQLPQSPVLSFKGGKN 331
DB 265 FVMTKTLNRLARDQTQPMCAFNDKVYATYQVTRGLASLNSC-----VDPILYFLAGDT 319

QY 332 -RVRLQKLRQNKLGHP 348
DB 320 FRRLSRATKRSRRSEP 337

RESULT 13
P2YR_MELGA STANDARD; PRT; 362 AA.
AC P49652;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DT P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (6H1 orphan receptor).
P2RY1.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94335907; PubMed=8058061;
RA Filtz T.M., Li Q., Boyer J.L., Nicholas R.A., Harden T.K.;
RT "Expression of a cloned P2Y purinergic receptor that couples to
RT phospholipase C."
RL Mol. Pharmacol. 46:8-14(1994).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=97382456; PubMed=9240460;
RA Li Q., Schachter J.B., Harden T.K., Nicholas R.A.;
RT "The 6H1 orphan receptor, claimed to be the p2y5 receptor, does not
RT mediate nucleotide-promoted second messenger responses."

```

```

RL Biochem. Biophys. Res. Commun. 236:455-460(1997).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
CC INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: MAINLY FOUND IN BLOOD, BRAIN, AND LUNG. TO A
CC LESSER EXTENT IN STOMACH, GUT, AND SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U09842; AAA18784.1; -.
CC DR EMBL; AF012103; AAB65428.1; -.
CC DR HSSP; P34996; 1DDD.
CC DR InterPro; IPR000276; GPCR_Rhodopsn.
CC DR Pfam; PF00001; 7tm1; 1.
CC DR PRINTS; PR00237; GPCRHOODPSN.
CC DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
CC DR PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 42 63 1 (POTENTIAL).
CC FT DOMAIN 64 76 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 77 98 2 (POTENTIAL).
CC FT DOMAIN 99 115 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 116 136 3 (POTENTIAL).
CC FT DOMAIN 137 155 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 156 177 4 (POTENTIAL).
CC FT DOMAIN 178 207 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 208 227 5 (POTENTIAL).
CC FT DOMAIN 228 254 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 255 274 6 (POTENTIAL).
CC FT DOMAIN 275 292 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 293 317 7 (POTENTIAL).
CC FT DOMAIN 318 362 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 113 191 BY SIMILARITY.
CC FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 362 AA; 41180 MW; 3E128AB9EB64349C CRC64;

Query Match 17.9%; Score 346; DB 1; Length 362;
Best Local Similarity 30.5%; Pred. No. 7.7e-16;
Matches 97; Conservative 49; Mismatches 128; Indels 44; Gaps 10;

QY 53 YLSLAYSIFILGLPLNGTVLWHSWGQTKRWSGQKRWSCATYLVNLMVADLLVYL-LPFLIITYS 111
DB 42 YLPTVIIVITGFLGNSVAIMWFHMRPWSGISVYMFNLADFLVLTLPALIFYF 101

QY 112 LDRWPFGECLCKLVHFLFYINLYGSILLTCTISVHOFGLGVCHPLCSLPYRTRRHAWLGT 171
DB 102 NKTDFWGDVCKLQRFIFHVNLYGSILLTCTISVHRVTGVVHPLKSLGRKKNAVYVS 161

QY 172 STTWALVQLLPTLAFSHDYNQMI-WYDWTSOENFDRFLFAYGIVLTL-----SGFLS 226
DB 162 SLVWALVAVIAPILFYSGTVGVRNKTTCYDTTADYLRSYFVSMCTTFVFCIPFIV 221

QY 227 LLGHGVLFTDQGEPDQARGEPEHEDRQSPSOVHPDPTGVVPLHPLFCALPYHSL-LLP 285
DB 222 ILGCVGLI-----VKALIYKDLNSPLRRK-----SIVLIIVLVFAVSYLP 264

QY 286 HLLSAFSGLPALD---GSQCGLQD-----MEASGECEQLPQSPVLSFKGGKN 331
DB 265 FVMTKTLNRLARDQTQPMCAFNDKVYATYQVTRGLASLNSC-----VDPILYFLAGDT 319

```

```
QY 332 -RVRLQKLQKNGEHP 348
| | | : | | :
Db 330 FRRRLSRATKRSRSEP 337

RESULT 14
P2YR_RAT
ID P2YR_RAT STANDARD; PRT; 373 AA.
AC P49651;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purine receptor)
GN P2Y1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Insulinoma;
RX MEDLINE=95298025; PubMed=7779087;
RA Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RT "Cloning of rat and mouse P2Y purinoceptors.";
RL Biochem. Biophys. Res. Commun. 211:211-218(1995).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
CC OF EXTRACELLULAR ATP ON INSULIN SECRETION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MUSCLE, HEART, LIVER, KIDNEY,
CC LUNG, BRAIN, SPLEEN, BUT NOT IN TESTIS.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U22830; AAA91303.1; -
DR HSSP; P34996; 1DD0.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52
FT TRANSMEM 53 74
FT DOMAIN 75 87
FT TRANSMEM 88 109
FT DOMAIN 110 126
FT TRANSMEM 127 147
FT DOMAIN 148 166
FT TRANSMEM 167 188
FT DOMAIN 189 218
FT TRANSMEM 219 238
FT DOMAIN 239 265
FT TRANSMEM 266 285
FT DOMAIN 286 303
FT TRANSMEM 304 328
FT DOMAIN 329 373
FT DISULFID 124 202
FT CARBOHYD 11 11
FT CARBOHYD 27 27
FT CARBOHYD 113 113
FT CARBOHYD 197 197
FT SEQUENCE 373 AA; 42321 MW; 6DDF676287B5E648 CRC64;
SQ
Query Match 17.0%; Score 329.5; DB 1; Length 373;
Best Local Similarity 36.7%; Pred. No. 9.2e-15;
```

```
Matches 77; Conservative 36; Mismatches 82; Indels 15; Gaps 4;
QY 53 YLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATYLVNLMVADLYVL-LPFLIITYS 111
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 53 YLPAVYIILVFIIGLGNVAIMFVHKPKWSGISVYMFNLCALADFLYVLTLPALIFYF 112
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 112 LDRWPFGEELCKLVHFLFYINLYGSIILLTCTISVHQFLGVCHPLCSLPYRTRHAWLGT 171
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 113 NKTDWIFGDVMCKLQRFIFHNLYGSIILFTCTISAHRYSGVVYPLKSLGRLKKNAIYVS 172
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 172 STTWALVVLQPLTLAFSHTDY-INGQMIWDMTQENFDRLFAYGIVTLTSGF----LS 226
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 173 VLWMLIVVVAISPILFYSGTGIRKNKTVTCYDSTDEYLSRYFIYSMCTVAMFCIPLVL 232
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 227 LLGHFGVLFTDQGPQDQARPEHEDROHSP 256
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 233 ILGCGIGLI-----VRALIYKDLNSP 253
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :

RESULT 15
P2YR_HUMAN
ID P2YR_HUMAN STANDARD; PRT; 373 AA.
AC P47900;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purine receptor).
GN P2Y1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RX MEDLINE=96257237; PubMed=8666290;
RA Leon C., Vial C., Cazenave J.-P., Gachet C.;
RT "Cloning and sequencing of a human cDNA encoding endothelial P2Y1
RT purinoceptor.";
RL Gene 171:295-297(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96158962; PubMed=8579591;
RA Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,
RA Kunapuli S.P.;
RT "Cloning and chromosomal localization of the human P2Y1
RT purinoceptor.";
RL Biochem. Biophys. Res. Commun. 218:783-788(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96205320; PubMed=8630005;
RA Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
RA Boeynaems J.M.;
RT "Cloning and tissue distribution of the human P2Y1 receptor.";
RL Biochem. Biophys. Res. Commun. 221:588-593(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=Blood;
RA Leon C., Vial C., Weber J., Cazenave J.-P., Gacher C.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
RX TISSUE=Platelet;
RX MEDLINE=98113162; PubMed=9442040;
RA Jin J., Daniel J.L., Kunapuli S.P.;
RT "Molecular basis for ADP-induced platelet activation. II. The P2Y1
RT receptor mediates ADP-induced intracellular calcium mobilization and
RT shape change in platelets.";
RL J. Biol. Chem. 273:2030-2034(1998).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. IN PLATELETS BINDING TO ADP LEADS TO MOBILIZATION OF
CC INTRACELLULAR CALCIUM IONS VIA ACTIVATION OF PHOSPHOLIPASE C, A
CC CHANGE IN PLATELET SHAPE, AND PROBABLY TO PLATELET AGGREGATION.
```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- INDUCTION: REPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS  
CC A3P5P, A3P5P AND A2P5P. THESE INHIBIT CALCIUM ION MOBILIZATION  
CC AND SHAPE CHANGE IN PLATELETS.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z49205; CAA89066.1; -  
DR EMBL; U42030; AAA97873.1; -  
DR EMBL; U42029; AAA97872.1; -  
DR EMBL; S81950; AAB47091.1; -  
DR EMBL; AJ006945; CAA07339.1; -  
DR EMBL; AF018284; AAB94556.1; -  
DR HSSP; P34996; 1DDD.  
DR Genew; HGNC:8539; P2RY1.  
DR MIM; 601167; -  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_FL1; 1.  
DR PROSITE; PS0262; G-PROTEIN\_RECEPTOR\_FL2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;  
KW Blood coagulation.  
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 53 74 1 (POTENTIAL).  
FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 88 109 2 (POTENTIAL).  
FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 127 147 3 (POTENTIAL).  
FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 167 188 4 (POTENTIAL).  
FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 219 238 5 (POTENTIAL).  
FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 266 285 6 (POTENTIAL).  
FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 304 328 7 (POTENTIAL).  
FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 124 202 BY SIMILARITY.  
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 136 138 MISSING (IN REF. 1).  
SQ SEQUENCE 373 AA; 42071 MW; 4DC7C668B4145392 CRC64;

Query Match 16.9%; Score 328; DB 1; Length 373;  
Best Local Similarity 29.9%; Pred No. 1.2e-14;  
Matches 93; Conservative 51; Mismatches 123; Indels 44; Gaps 10;

QY 53 YLSLAYSIIFILGLPLNGTVLHSHWGQTRKWSGATYLVNLMVADLYVL-LPFLIITYS 111  
DB 53 YLPVAVILVFIIGFLGNSVAINMFVFMKRPWSGISVYMFNLALADFLVLTLPALIFYF 112  
QY 112 LDDRPFGECLKLVHFLYINLYGSIILLTCISVHQFGLVCHPLCSLPYRTRRHAWLGT 171  
DB 113 NKTWDWIFGDAMCKLQRFIFHVNLYGSIILFTCSIAHRYSGVYVPLKSLGRLKKNAICIS 172  
QY 172 STTWALVWLQLPLTFLAFSTHYDINGOMI-WYDMTQENFDRLFAFGVLTLSGF----LS 226  
DB 173 VLWLVVVAISILFYSGTVKKNKTIICYDITSDEYLSRFSIYSMCTTVMFCVPLVL 232  
QY 227 LIGHGVLFDTGQEPDQARGEHPHQSPSQVHPDHTGVPWPLHPLFCALPYHSL-LLP 285  
DB 233 ILGCYGLI-----VRALIVKLDLNSPLRK-----SIVLVIIVLTVFAVSYP 275

QY 286 HHLLSAFSGLPALD---GSQCGLQD-----MEASGECEQLPQSPVLSFKGGKN 331  
DB 276 FHVMTMNLRLARLDFTPAMCAFNDRVYATYQVTRGLASLNSC-----VDPILYFLAGDT 330  
QY 332 -RVRLLOKLRQ 341  
DB 331 FRRRLSRATRK 341  
Search completed: May 22, 2003, 12:50:04  
Job time : 16 secs

Result No.	Query <sup>8</sup>			DB	ID	Description
	Match	Length	Score			
1	421	21.7	374	13	Q57466	O57466 meleagris g
2	395	20.4	328	11	Q9ERK9	Q9ERK9 mus musculus
3	359	18.5	357	13	Q9DE05	Q9DE05 raja erinac
4	329.5	17.0	361	13	Q90X57	Q90X57 xenopus lae
5	278.5	14.4	337	4	Q96P68	Q96P68 homo sapien
6	249	12.9	399	11	Q8R311	Q8R311 mus musculus
7	248.5	12.8	359	6	Q9N0U1	Q9N0U1 ovis aries
8	248	12.8	315	6	Q9GKR7	Q9GKR7 sus scrofa
9	244.5	12.6	317	11	Q99MT6	Q99MT6 mus musculus
10	242	12.5	359	6	Q9GLN9	Q9GLN9 pan troglod
11	241.5	12.5	359	4	Q8TBK4	Q8TBK4 homo sapien
12	235	12.1	358	13	Q9DGT1	Q9DGT1 gallus galli
13	233	12.0	359	11	Q9EQR9	Q9EQR9 meriones un
14	232.5	12.0	330	4	Q9BXA5	Q9BXA5 homo sapien
15	231.5	12.0	334	4	Q8TDQ8	Q8TDQ8 homo sapien
16	231.5	12.0	423	5	Q964D4	Q964D4 periplaneta







Db	62	CGNSWIIYILRYAKMKTAFTNFIILNLAHADELMLSVPL- VTSTLLRHWPFGALLCRL	120
Qy	126	VHELFYNLGSLLLTCTISVHGFGLGVCHPLCSLPYRTRRRHAWLGSTTWAALVVLLOLPT	185
Db	121	VLSVDVNMFTSIYCLTVLSVDRYVAVHPKARVRRYTKVKNVLSLVLILPI	180
Qy	186	LAFSHTDYINGQMIWDMTSEQNFDR-LFAYGIVLTLGSLGSLG	229
Db	181	VWFSRTAANSDGTACVACNMLMPEAQRLVGVFLVLTFLMGFLLPVG	225
RESULT 9			
ID	Q99MT6	PRELIMINARY;	PRT; 317 AA.
AC	Q99MT6;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	G-protein coupled receptor GPR91.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
ON	NCBI_TaxID=10090;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL;		
RX	MEDLINE=21172992; PubMed=11273702;		
RT	Wittenberger T., Schaller H.C., Hellebrand S.;		
RT	"An expressed sequence tag (est) data mining strategy succeeding in		
RL	the discovery of new G-protein coupled receptors.,"		
RL	J. Mol. Biol. 307:799-813(2001).		
DR	EMBL; AF293367; AK01867.1; -.		
DR	HSSP; P34996; IDDD.		
DR	MGD; MGI:1934135; Gpr91.		
DR	InterPro; IPR000276; GPCR_Rhodpsn.		
DR	Pfam; PF00001; 7tm1; 1.		
DR	PRINTS; PR00237; GPCRHHODOPS.		
DR	PROSITE; PS00237; G_PROTEIN_RECP_FL1; UNKNOWN_1.		
DR	PROSITE; PS0262; G_PROTEIN_RECP_FL2; 1.		
DR	Receptor.		
SW	SEQUENCE 317 AA; 36701 MW; 4E22F0608F928E36 CRC64;		
Query Match			
Best Local Similarity 12.6%; Score 244.5; DB 11; Length 317;			
Matches 88; Conservative 50; Mismatches 135; Indels 37; Gaps			
Qy	48	KYQVYLSAYSIIIFILGPLNGTVLWHSNGQTKRSCATTYLVNLMVADLLXV-LLPEL	106
Db	19	KY---YLSAFYAIEFIPLGNVTYVFGVLFQCKMKNSSNVYLFNLSISDFAPLCLPLIL	75
Qy	107	IITYLDDRRWPFCELLKLVHFLFYINLVXGSIILLTCTISVHGFGLGVCHPLCSLPYRTRRH	166
Db	76	IKSYA-NDKTYGDVLCISNRYVHLNLTSMLLTVISMDRVLLMKYFREFHFLOKKEE	134
Qy	167	AWLGSTTWAALVVLLOLPTLAFSHTDYINGQMIWDMTSEQNFDRLFAYGIVLTLGSGF--	224
Db	135	ALLISLAVNALVTLVPLMLTFTNSVPKEGSCIDYASSGNPEHNLIIYSLCTLLGLFI	194
Qy	225	-LSLLGHF---GVLFDTGGQEPDQARCEPHEDRQHSQVHPDHTGVWPLHPLFCALPHY	280
Db	195	PLSVMGFFYYKMWVFLKRRSQOQATALPLDKPQRL-----VVLAVVIFS	238
Qy	281	SLPLPHILLSAFSGLPALDQSCGLQDMEASGCEQLPOP-----SPVLISF-KGGKN	331
Db	239	ILFTPYHIMNLRIASRLDSWPGQC-TQKAISYITLTPLAFLNSAINPIFFVLMDGHI	297
Qy	332	RVRLQKLKQ 341	
Db	298	REMLISKFRQ 307	



9



Db 138 FAILISLAIWVLTLELPILPLINPVITDNGTTCNDFASSGDPNYNLIYSMCITLIGFL 197

Search completed: May 22, 2003, 12:50:46  
Job time : 36 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 21:29:43 ; Search time 304 Seconds  
(without alignments)  
2666.840 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILLPSRSGSRGRL.....QNKLGHPAGKRCPLNRS 360

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+...  
-DB=US10023586...  
-LOOPEXT=0  
-LIST=45  
-MODE=LOCAL  
-USER=US10023586...  
-NO\_XLPXY  
-WARN\_TIMEOUT=30  
-YGAPOP=10

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1936	100.0	3143	24	Human genomic clon
2	1898	98.0	1063	24	Human DNA for pote
3	1798	92.9	1002	22	Human G protein-co
4	1785	92.2	1076	22	Human G-protein co
5	1541	79.6	850	24	Human P2Y-like G p
6	921	47.6	510	22	Human ngPCR12 codi
7	732	37.8	681	22	Human CDNA encodin
8	681	35.2	585	22	Human purinergic r
9	453	23.4	1163	22	Human P2Y recepto
10	412	22.8	1429	18	Human purinergic r
11	420	21.7	984	17	Mouse pancreas G-p
12	418	21.5	1842	16	Human P20 receptor
13	415	21.4	2138	24	Mus musculus GPCR
14	415	21.4	2138	24	Mus musculus GPCR
15	413	21.3	984	17	Human placenta G-p
16	402	20.8	984	18	P2U purinergic rec
17	402	20.8	984	18	Human placenta pur
18	370	19.1	551	22	Human G protein-co
19	328	16.9	6721	24	Purinergic recepto
20	327	16.9	6721	24	Purinergic recepto
21	294	15.2	450	21	Human secreted exp
22	292	15.1	234	22	Human CDNA encodin
23	281	14.5	1020	22	Human ngPCR54 codi
24	278	14.4	1011	24	Human G protein-co
25	278	14.4	1011	24	Human CDNA encodin
26	278	14.4	1011	24	Human G-protein co
27	278	14.4	1014	24	Human P2Y-like rec
28	278	14.4	1014	24	Human AXOR89 (G-pr
29	278	14.4	1014	24	Human purinergic-r
30	278	14.4	1014	24	Human G-protein co
31	278	14.4	1014	24	Human P2Y1-li enco
32	278	14.4	1288	24	Human CDNA encodin
33	278	14.4	1729	22	Human prostate exp
34	278	14.4	1729	23	Human prostate exp
35	278	14.4	1729	23	Human prostate exp
36	278	14.4	1729	23	Human prostate exp
37	278	14.4	1729	23	Human prostate exp
38	278	14.4	1729	23	Human prostate exp
39	278	14.4	5435	24	Human purinergic-r
40	278	14.4	9905	24	Human polynucleoti
41	278	14.4	1313	22	Human G-protein co
42	275	14.2	1301	17	Human G-protein co
43	275	14.2	1301	21	Human G-protein co
44	275	14.2	1301	21	Human G-protein co
45	269	13.9	2162	19	Human mu-opioid re

ALIGNMENTS

RESULT 1

AA517746

ID AAS17746 standard; DNA; 3143 BP.

XX AAS17746;

AC AAS17746;

XX 26-FEB-2002 (first entry)

XX Human genomic clone for P2Y-like G protein-coupled receptor..

DE Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;

XX Chronic obstructive pulmonary disease; nervous system disease;

KW Parkinson's disease; multiple sclerosis; dementia; stroke;

KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;

KW bacterial infection; fungal infection; protozoan infection;

KW viral infection; pain; cancer; anorexia; bulimia; asthma;

KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;

KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;



psychotic disorder; neurological disorder; anxiety; schizophrenia;  
manic depression; delirium; severe mental retardation; dyskinesia.

Homo sapiens.

Key Location/Qualifiers  
CDS 520..2498

/\*tag= a  
/product= "P2Y-like GPCR"

W0200185764-A2.

15-NOV-2001.

09-MAY-2001; 2001WO-EP05244.

11-MAY-2000; 2000US-203582P.

21-FEB-2001; 2001US-269857P.

(FARB ) BAYER AG.

Ramakrishnan S;

WPI: 2002-075242/10.

P-PSDB; AAU11251.

New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders

Disclosure, Fig 1; 114pp; English.

The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR.

In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, osteoporosis, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies, psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic depression, delirium, severe mental retardation or dyskinesias). The present sequence is a genomic clone encoding the P2Y-like GPCR of the invention.

Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 other;

#### Alignment Scores:

Pred. No.:	1-29e-174	Length:	3143
Score:	1936.00	Matches:	360
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-023-586B-2 (1-360) x AAS17746 (1-3143)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerArgSerArgGlyAlaLeu 20  
DB 436 ATGCTGTCCATTTCCTTCACGGGGGAGCAGAGGGGAGCGCTCGTGAGCTCTG 495

QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluIn 40  
DB 496 CTCCTGGAGGAGGCTCCCGGACATGGAGAGGTGGACATGATACATCAGAGCAAA 555  
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
DB 556 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 615  
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
DB 616 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGGCCAAACC 675  
QY 81 LysArgTyrSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
DB 676 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCGACCTGCTTAT 735  
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120  
DB 736 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 795  
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
DB 796 CTGCTCTGCAAGCTGGTGACCTCTCTGTTCTATATCACTTTTACGGCAGCATCTCTGCTG 855  
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
DB 856 CTGACCTGATCTCTCTGTCACCACTTCTAGGTGTCTGCCACCACTGTTGCTGCTGCC 915  
QY 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180  
DB 916 TACCGGACCGCAGGATGCTGCTGGCCACCACTGCGCCCTGGTGGTGGTCTC 975  
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
DB 976 CAGCTGCTGCCACACTGGCTTCTCCACACGAGCTACATCAATGGCCAGAGATGATCTGG 1035  
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
DB 1036 TATGACATGACACCAAGAGAAATTTGATCGCTTTTGCCTACGGCATAGTTCTTGACA 1095  
QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
DB 1096 TTGTCTGGCTTCTTCTCCCTCTGGTCAATTTTGGTGTCTATTTCACTGATGGTCAAGAG 1155  
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
DB 1156 CCGTATCAAGCCAGAGAGAACCTCATGAGACAGGCAACACACCCCGGCGGAGTCCAT 1215  
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
DB 1216 CCGGACCATCTACTGCTGTGGCTTTCACCTCTGTTTGTGGCTTCCATATCATC 1275  
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
DB 1276 TCGCTCTCTTACCTCACCATCTGCTTCTCTTCAGAGCTGCCAGCTCTTGATGGC 1335  
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320  
DB 1336 AGCCAGTGTGGCTTACAGATATGGAGCCCTCTGGTGTGTGTGAGCAGCTGCCTCAACCC 1395  
QY 321 SerProValLeuSerPheLeuGlyLysValAsnArgValArgLeuLeuGlnLysLeuArg 340  
DB 1396 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAAATGAGG 1455  
QY 341 GlnAsnLysLeuGlyGluHisProAlaClyArgLysArgCysProGlyLeuAsnArgSer 360  
DB 1456 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACATCT 1515

RESULT 2

AAS98135

ID AAS98135 standard; DNA; 1063 BP.

XX

AC AAS98135;





616 ATCTTTATCTAGGGTGGCCACTAAATAGGCACTGCTCTGTGGCACTCCTGGGGCCAAACC 675  
QY LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
Db 676 AAGCGCTGGAGCTGGCCACCACCTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 735  
QY 101 ValLeuLeuProPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
Db 736 GTGCTATTGGCTTCTCCATCATCATCTACTACTACTAGTAGACAGGTGGCCCTCGGGGAG 795  
QY 121 LeuLeuLeuLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
Db 796 CTGCTGTCGACAGCTGGTGCATCTCTGTCATATACACCTTACGGCAGCATCTGCTG 855  
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
Db 856 CTGACCTGCATCTCTGTGCACAGTCTCTAGTGTGTGGCCACCATCTGTTCGTGCTGCC 915  
QY 161 TyrArgThrArgGlyHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180  
Db 916 TACCGGACCCGACGACATGCTGGTGGCGCACCATCCACCATCTGGTGGCTGCTCTC 975  
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
Db 976 CAGCTGCTGCCACACTGCTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 1035  
QY 201 TyrAspMetThrSerGlnGlnAsnPheAspArgLeuPheAlaTyrGlyLeuValLeuThr 220  
Db 1036 TATGACATGACGACGACAGAGAAATTTGATCGCTTTTGGCTACGGCATAGTCTGACA 1095  
QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
Db 1096 TTGCTGGCTTTCTTCCCTCTGCTGCTATTTGGTGTCTATTCTACTGATGTCAGGAG 1155  
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
Db 1156 CCGTATCAGCCAGAGAGAACTCTATGAGGACACGACACGACCCGACGAGGTCAT 1215  
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
Db 1216 CCGGACCATCTACTGGTGTGGCTCTTCCACCTCTGTTTGTGCCCTTCCATATCAC 1275  
QY 281 SerLeuLeuLeuProHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
Db 1276 TCCTCTCTTACTCACCATCTGCTTTCTGCTTTCTCAGACTCCAGCTCTGTGAGGC 1335  
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320  
Db 1336 AGCCAGTGTGGCTACAGATATGAGGCTCTGGTGTGAGTGTGAGCAGCGCTCAACCC 1395  
QY 321 SerProValLeuSerPheLeuGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340  
Db 1396 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAAATGAGG 1455  
QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
Db 1456 CAGAACAAAGTTGGTGTGATCCAGCTCCAGTGGGAGGAGATGCCCGAGGTGTGAACAGATCT 1515

RESULT 4  
AC025960/c AC025960 164502 bp DNA linear HTG 13-APR-2000  
DEFINITION Homo sapiens chromosome 3 clone RP11-170K4 map 3, WORKING DRAFT  
SEQUENCE, 15 unordered pieces.  
AC025960  
VERSION AC025960.2 GI:7547202  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 164502)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 3, clone RP11-170K4

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 164502)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campoliano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,  
Dodgson, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galgan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehocsky, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Rogov, P., Rothman, D.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Spencer, B.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Stojanovic, N., Subramanian, A., Talamas, J.,  
Stange-Thomann, N., Tirrell, A., Travers, M., Trigilio, J.,  
Tesfaye, S., Theodore, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 13, 2000 this sequence version replaced gi:7328839.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L8491  
Center clone name: 170\_K\_4  
----- Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 156927 bases at least Q40  
Consensus quality: 160376 bases at least Q30  
Consensus quality: 161800 bases at least Q20  
Insert size: 170000; agarose-fp  
Insert size: 163102; sum-of-contigs  
Quality coverage: 4.4 in Q20 bases; agarose-fp  
Quality coverage: 4.5 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
\* 2918 3017: gap of 100 bp  
\* 3018 6310: contig of 3293 bp in length  
\* 6311 6410: gap of 100 bp  
\* 6411 9447: contig of 3037 bp in length  
\* 9448 9547: gap of 100 bp  
\* 9548 14470: contig of 4923 bp in length  
\* 14471 14570: gap of 100 bp  
\* 14571 20528: contig of 5958 bp in length  
\* 20529 20628: gap of 100 bp  
\* 20629 24667: contig of 4039 bp in length  
\* 24668 24767: gap of 100 bp  
\* 24768 29367: contig of 4600 bp in length  
\* 29368 29467: gap of 100 bp  
\* 29468 36237: contig of 6770 bp in length







```

repeat_region      /rpt_family="AT_rich"
7595..7717
/rpt_family="MERSA"
complement(7788..8198)
/rpt_family="MLTII"
complement(8536..8591)
/rpt_family="L2"
9634..9936
/rpt_family="AluSx"
complement(10766..10814)
/rpt_family="MIR"
complement(10848..10964)
/rpt_family="MIR"
17399..17431
/rpt_family="TTAAA"n"
complement(17432..17546)
/rpt_family="MIR"
17637..18073
/rpt_family="Charlieb"
18278..18313
/rpt_family="(TG)n"
21428..21841
/rpt_family="MSTD"
complement(21847..21926)
/rpt_family="LTR33A"
22387..22631
/rpt_family="MLTIG1"
complement(22768..23212)
/rpt_family="Tigger3(Golem)"
complement(23222..23279)
/rpt_family="Tigger3(Golem)"
23311..23437
/rpt_family="MLTIE2"
24105..24225
/rpt_family="MIR"
24226..24517
/rpt_family="MER33"
24518..24629
/rpt_family="MIR"
complement(24639..24724)
/rpt_family="L2"
25420..25994
/rpt_family="L2"
26024..26059
/rpt_family="(TG)n"
26565..27405
/rpt_family="LIMB8"
complement(27789..27855)
/rpt_family="L2"
28628..28778
/rpt_family="MER77"

Alignment Scores:
Pred. No.:      2.75e-163      Length:      167084
Score:          1936.00      Matches:      360
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              9              Gaps:      0

US-10-023-586B-2 (1-360) x AC092999 (1-167084)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
Db 160921 ATGCTGTCCATTTCTCTCCAGGGAAGCAGAGAGCGGCGCTGGAGCTCTG 160980

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
Db 160981 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATCATCATCAGGAACAA 161040

Qy 41 GlyLeuGluGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 161041 GGTCTGTCCAGTTCTCAGAGAAGTACAAGCAGAGTCTACCTCTCCCTGGCCTACAGTATC 161100

```

```

Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr 80
Db 161101 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGCCAAACC 161160

Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 161161 AAGCGTGGAGCTGTGCCACCACCTATCTGTGAACCTGTGTGGCGGACCTGCTTTAT 161220

Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTyrProPheGlyGlu 120
Db 161221 GGGTATTTGGCTTCTCTCATCTCACCTACCTACCTAGATGACAGGTGGCCCTTCGGGGAG 161280

Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 161281 CTGCTCTCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTGTG 161340

Qy 141 LeuThrCysLysLeuSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 161341 CTGACCTGCATCTGTGTGCACCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161400

Qy 161 TyrArgThrArgArgHisAlaTyrLeuGlyThrSerThrThrTyrAlaLeuValValLeu 180
Db 161401 TACCGGACCCGAGGCACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161460

Qy 181 GlnLeuLeuProThrLeuAlaGlySerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
Db 161461 CAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161520

Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 161521 TATGACATGACCAAGCAAGAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161580

Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 161581 TTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161640

Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 161641 CCTGTATCAAGCCAGAGAGAACTCTCATGAGGACAGCAACACAGCCGAGCAGGTCCAT 161700

Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheValAlaLeuProTyrHis 280
Db 161701 CCGGACCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161760

Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db 161761 TCGCTCTCTCTACCTCACCATCTGTCTTCTGTCTTCTCTCAGGAGTGCAGCTTTGATGC 161820

Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
Db 161821 AGCCAGTGTGGCTACAAGATATGGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161880

Qy 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuAla 340
Db 161881 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAAACCTGAGG 161940

Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db 161941 CAGAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGACACATCT 162080

RESULT 6
AC021773      163958 bp      DNA      linear      HTG 10-SEP-2000
LOCUS      Homo sapiens clone RP11-12N17, WORKING DRAFT SEQUENCE, 15 unordered
DEFINITION      pieces.
ACCESSION      AC021773
VERSION      AC021773.6 GI:10047750
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



# REFERENCE AUTHORS JOURNAL REFERENCE AUTHORS

1 (bases 1 to 163958)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome, clone RP11-12N17  
 Unpublished  
 2 (bases 1 to 163958)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Becker, R., Bede, F.,  
 Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,  
 Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,  
 Ferraira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
 Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
 MacDonald, P., Marquis, J., McEwan, P., McGurk, A., McKernan, K.,  
 McPheders, R., Melidrim, J., Meneus, L., Morrow, J., Naylor, J.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
 Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 10, 2000 this sequence version replaced gi:7248967.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L3302  
 Center clone name: 12\_N\_17  
 ----- Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 152766 bases at least Q40  
 Consensus quality: 158075 bases at least Q30  
 Consensus quality: 160495 bases at least Q20  
 Insert size: 174000; agarose-fp  
 Insert size: 162558; sum-of-contigs  
 Quality coverage: 4.5 in Q20 bases; agarose-fp  
 Quality coverage: 4.9 in Q20 bases; sum-of-contigs

## TITLE JOURNAL COMMENT

NOTE: This is a 'working draft' sequence. It currently  
 consists of 15 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 7254: contig of 7254 bp in length  
 7255 7354: gap of 100 bp  
 7355 8612: contig of 1258 bp in length  
 8613 8712: gap of 100 bp  
 8713 11463: contig of 2751 bp in length  
 11464 11563: gap of 100 bp  
 11564 14507: contig of 2944 bp in length  
 14508 14607: gap of 100 bp  
 14608 22452: contig of 7845 bp in length  
 22453 22552: gap of 100 bp  
 22553 35086: contig of 12534 bp in length  
 35087 35186: gap of 100 bp  
 35187 44278: contig of 9092 bp in length  
 44279 44378: gap of 100 bp  
 44379 56786: contig of 12408 bp in length  
 56787 56886: gap of 100 bp

\* 56887 65270: contig of 8384 bp in length  
 \* 65271 65370: gap of 100 bp  
 \* 65371 76394: contig of 11024 bp in length  
 \* 76395 76494: gap of 100 bp  
 \* 76495 100320: contig of 23826 bp in length  
 \* 100321 100420: gap of 100 bp  
 \* 100421 113127: contig of 12707 bp in length  
 \* 113128 113227: gap of 100 bp  
 \* 113228 126756: contig of 13529 bp in length  
 \* 126757 126856: gap of 100 bp  
 \* 126857 151061: contig of 24205 bp in length  
 \* 151062 151161: gap of 100 bp  
 \* 151162 163958: contig of 12797 bp in length.

## FEATURES

source  
 1. 163958  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-12N17"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1. 7254  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 7355. 8612  
 /note="assembly\_fragment"  
 8713. 11463  
 /note="assembly\_fragment"  
 11564. 14507  
 /note="assembly\_fragment"  
 14608. 22452  
 /note="assembly\_fragment"  
 22553. 35086  
 /note="assembly\_fragment"  
 35187. 44278  
 /note="assembly\_fragment"  
 44379. 56786  
 /note="assembly\_fragment"  
 56887. 65270  
 /note="assembly\_fragment"  
 65371. 76394  
 /note="assembly\_fragment"  
 76495. 100320  
 /note="assembly\_fragment"  
 100421. 113127  
 /note="assembly\_fragment"  
 113228. 126756  
 /note="assembly\_fragment"  
 126857. 151061  
 /note="assembly\_fragment"  
 151162. 163958  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right

BASE COUNT 49230 a 32369 c 32422 g 48536 t 1401 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4,05e-162 Length: 163958  
 Score: 1923.00 Matches: 360  
 Percent Similarity: 99.72% 100% Conservative: 0  
 Best Local Similarity: 99.72% Mismatches: 1  
 Query Match: 99.33% Indels: 1  
 DB: 2 Gaps: 0

US-10-023-586B-2 (1-360) x AC021773 (1-163958)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
 Db 98103 ATGCTGTCATTTGCTTCCTCCAGGGGAAACAGACGGGAGCGCTGCTG 98162  
 Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
 Db 98163 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGATACATCAGGACACAA 98222





AF411108	AF411108.1	GI:16566321	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Lee, D.K., Nguyen, T., Lynch, K.R., Cheng, R., Vanti, W.B., Arkhitko, O., Lewis, T., Evans, J.F., George, S.R. and O'Dowd, B.F.		
TITLE	Discovery and mapping of ten novel G protein-coupled receptor genes		
JOURNAL	Gene 275 (1), 83-91 (2001)		
MEDLINE	21458557		
PUBMED	11574155		
REFERENCE	2 (bases 1 to 1019)		
AUTHORS	Lee, D.K., Nguyen, T., Lynch, K.R., Cheng, R., Vanti, W.B., Arkhitko, O., Lewis, T., Evans, J.F., George, S.R. and O'Dowd, B.F.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-AUG-2001) Department of Pharmacology, University of Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada		
FEATURES	Location/Qualifiers		
source	1..1019		
gene	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	1..1019		
	/gene="GPR79"		
	/pseudo		
BASE COUNT	200 a 308 c 252 g 259 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,0e-154	Length:	1019
Score:	1806.00	Matches:	339
Percent Similarity:	99.71%	Conservative:	0
Best Local Similarity:	99.71%	Mismatches:	0
Query Match:	93.29%	Indels:	1
DB:	9	Gaps:	0
US-10-023-586B-2 (1-360) x AF411108 (1-1019)			
Qy	1 MetLeuSerIleLeuLeuProSerATGGLYSerArgSerGlySerArgArgGlyAlaLeu	20	
Db	1 ATGCTGTCCATTTCCTTCCTTCAGGGAAGACGGAGCCGCTGCTGGAGCTCTG	60	
Qy	21 LeuLeuGluGlyAlaSerArgAspMetGLYSValAspMetAsnThrSerGlnGluGln	40	
Db	61 CTCCTGGAGGAGGCTCCCGGACATGGAGAGGTGCACATGAATACATCACAGACAA	120	
Qy	41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle	60	
Db	121 GGTCCTGCCAGTTCTCAGAGAAGTACAAGCAAGTTACCTCTCCCTGGCCTACAGTATC	180	
Qy	61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpPHisSerTrpGlyGlnThr	80	
Db	181 ATCTTTATCTAGGCTGCGCCACCTAAATGGCACGTCTGTGGCACTCCTGGGGCCAAACC	240	
Qy	81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr	100	
Db	241 AAGCGCTGGAGCTGCGCACCACTATCTGGTGAACCTGATGTTGGCGGACCTGCTTTAT	300	
Qy	101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGlu	120	
Db	301 GTGCTATTGGCTTCCTCATCATCACCTACTACTAGATCACAGGTGGCTCTCGGGGAG	360	
Qy	121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu	140	
Db	361 CTGCTCTGCAAGCTGGTGCACCTTCCTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTG	420	
Qy	141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro	160	
Db	421 CTGACCTGCATCTCTGTGCACCAAGTCTCCTAGGTGTGGCCACCACCTGTGCTGCTGCC	480	
Qy	161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeu	180	











## RESULT 5





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 20:51:54 ; Search time 3164 Seconds  
(without alignments)  
9952.350 Million cell updates/sec

Title: US-10-023-586B-1  
Perfect score: 1082  
Sequence: 1 atgtgtccattttgtctcc.....ccagggtgaacagatctgg 1082

oring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_hgt:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_hgt\_hum:\*
- 31: em\_hgt\_inv:\*
- 32: em\_hgt\_other:\*
- 33: em\_hgt\_mus:\*
- 34: em\_hgt\_pln:\*
- 35: em\_hgt\_rtd:\*
- 36: em\_hgt\_mam:\*
- 37: em\_hgt\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hgt\_hum:\*
- 40: em\_hgt\_mus:\*
- 41: em\_hgt\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1082	100.0	1082	6	AX458306	Sequence
2	1082	100.0	1334	9	AB065670	Homo sapi
3	1082	100.0	3143	6	AX365511	Sequence
c	1082	100.0	164502	2	AC026960	Homo sapi
4	1082	100.0	167084	9	AC092999	Homo sapi
5	1082	100.0	167084	9	AC092999	Homo sapi
6	1071	99.0	163958	2	AC021773	Homo sapi
7	1007	93.1	1019	9	AF411108	Homo sapi
8	1002.8	92.7	1020	6	AX458308	Sequence
9	1002.8	92.7	1076	6	AX375230	Sequence
10	850	78.6	850	6	AX365514	Sequence
11	649.4	60.0	209273	2	AC121579	Mus muscu
c	644.4	59.6	91107	2	AC116265	Rattus no
13	510	47.1	510	6	AX147768	Sequence
14	481.6	44.5	192524	2	AC120742	Rattus no
c	396.4	36.6	681	6	AX244718	Sequence
16	370	34.2	585	6	AX244727	Sequence
17	235.8	21.8	1163	5	GDP2Y3	X98283 G.domesticu
18	227.8	21.1	1074	5	AF069555	Meleagris
19	227.8	21.1	2025	9	HSU07225	Homo sapien
20	226.2	20.9	2257	9	BC028135	Homo sapi
21	224.6	20.8	2681	9	BC012104	Homo sapi
22	224.6	20.8	123185	2	AP000587	Homo sapi
23	224.6	20.8	196988	9	AP002761	Homo sapi
c	222	20.5	551	6	AX230145	Sequence
25	203.8	18.8	1842	6	I34573	Sequence 1
26	203.8	18.8	1842	6	I36524	Sequence 1
27	203.8	18.8	1842	6	I76280	Sequence 1
28	202.2	18.7	2138	6	AX454153	Sequence
29	202.2	18.7	2562	10	MUSP20REC	Sequence
30	202.2	18.7	2562	10	BC006613	Sequence
31	201.4	18.6	1977	5	AF031897	Mouse ATP r
32	200.6	18.5	175947	2	AC079641	Mus muscu
33	199	18.4	1131	10	S83099	P2U recepto
34	187.8	17.4	1868	10	RNU56839	Rattus norv
35	187.8	17.4	1918	10	RNU09402	Rattus norv
36	187.8	17.4	2011	10	RATP2U	Rattus norv
37	178.6	16.5	1232	9	HSGPCRE	H.sapiens g
38	178.6	16.5	1429	9	A62985	Sequence 1
39	178.6	16.5	1429	9	HSP2YA	H.sapiens p
c	178.6	16.5	178868	9	AL357752	Human DNA
40	177	16.4	178861	2	AC068231	Homo sapi
42	173.8	16.1	1651	9	HSU40223	Human uridi
43	170	15.7	234	6	AX244726	Sequence
44	165.2	15.3	1177	10	RNP244REC	R.norvegicu
45	164.8	15.2	984	6	AR109633	Sequence

ALIGNMENTS

RESULT 1	AX458306	Sequence 1	1082 bp	DNA	linear	PAT 08-JUL-2002
AX458306	LOCUS	Sequence 1 from Patent EP1215214.				
DEFINITION	AX458306	AX458306				
ACCESSION	AX458306.1	GI:21725001				
VERSION	human.					
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1					
TITLE	Fidock,M.D.					
JOURNAL	Novel polypeptide					
	Patent: EP 1215214-A 1 19-JUN-2002;					



Db 177 CTCCTGGAGGAGCCTCCGGACATGAGAGGTGGACATGAATACATCACAGGAACAA 236  
Qy 121 GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180  
Db 237 GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 296  
Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGTGTGGCACTCTCTGGGGCCAAACC 240  
Db 297 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGTGTGGCACTCTCTGGGGCCAAACC 356  
Qy 241 AAGCCTGGAGCTGTGCCACCACTATCTGTGGAACCTGTAGTGGTGGCCGACCTGCTTTAT 300  
Db 357 AAGCCTGGAGCTGTGCCACCACTATCTGTGGAACCTGTAGTGGTGGCCGACCTGCTTTAT 416  
Qy 301 GTGCTATTGGCCCTCTCTCATCATCACTACCTACTACCTAGATGACAGTGGCCCTTCGGGGAG 360  
Db 417 GTGCTATTGGCCCTCTCTCATCATCACTACCTACTACCTAGATGACAGTGGCCCTTCGGGGAG 476  
Qy 361 GTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTACGGCAGCATCTCGCTG 420  
Db 477 GTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTACGGCAGCATCTCGCTG 536  
Qy 421 CTGACCTGCATCTCTGTGCACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 480  
Db 537 CTGACCTGCATCTCTGTGCACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 596  
Qy 481 TACCGGACCCGAGCAGTCTCTGGCTGGCGCACAGCACACCTGGCGCCCTGTGTGCTCTC 540  
Db 597 TACCGGACCCGAGCAGTCTCTGGCTGGCGCACAGCACACCTGGCGCCCTGTGTGCTCTC 656  
Qy 541 CAGCTGTCTGCCACACTGGCTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600  
Db 657 CAGCTGTCTGCCACACTGGCTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 716  
Qy 601 TATGACATGACGAGCAGAGAAATTTGATCGGCTTTTGCCTACGGCATAGTTCTTGACA 660  
Db 717 TATGACATGACGAGCAGAGAAATTTGATCGGCTTTTGCCTACGGCATAGTTCTTGACA 776  
Qy 661 TTGCTCGGCTTTCTTCTCCCTCTCTGGTCAATTTTGGTGTGCTATTCACTGATGGTCAGGAG 720  
Db 777 TTGCTCGGCTTTCTTCTCCCTCTCTGGTCAATTTTGGTGTGCTATTCACTGATGGTCAGGAG 836  
Qy 721 CCTGATCAAGCCAGAGAGAACTCATGAGGACAGCCAAACAGCCCGAGCAGGTCCAT 780  
Db 837 CCTGATCAAGCCAGAGAGAACTCATGAGGACAGCCAAACAGCCCGAGCAGGTCCAT 896  
Qy 781 CCGGACCATCTACTGGTGTGGCTCTTCAACCTCTGTTTGTGGCCCTTCCATATCAC 840  
Db 897 CCGGACCATCTACTGGTGTGGCTCTTCAACCTCTGTTTGTGGCCCTTCCATATCAC 956  
Qy 841 TCGCTCTCTTACCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGCG 900  
Db 957 TCGCTCTCTTACCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGCG 1016  
Qy 901 AGCCAGTGTGGCTTACAAGATATGAGGCCCTCTGGTGTGATGTGAGCAGTGCCTCAACCC 960  
Db 1017 AGCCAGTGTGGCTTACAAGATATGAGGCCCTCTGGTGTGATGTGAGCAGTGCCTCAACCC 1076  
Qy 961 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTGAGG 1020  
Db 1077 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTGAGG 1136  
Qy 1021 CAGAAACAGTTGGGTGACATCCAGCTGGGAGGAGAGATGCCAGGCTTGAACAGATCT 1080  
Db 1137 CAGAAACAGTTGGGTGACATCCAGCTGGGAGGAGAGATGCCAGGCTTGAACAGATCT 1196  
Qy 1081 GG 1082  
Db 1197 GG 1198

AX365511 3143 bp DNA linear PAT 21-FEB-2002  
LOCUS Sequence 1 from Patent WO0185764.  
DEFINITION AX365511  
ACCESSION AX365511  
VERSION AX365511.1 GI:18873666  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Ramakrishnan,S.  
TITLE Regulation of human p2y-like g protein-coupled receptor  
JOURNAL Patent: WO 0185764-A 1 15-NOV-2001;  
Bayer Aktiengesellschaft (DE)  
FEATURES  
source Location/Qualifiers  
1..3143  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
520..>1374  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD24312.1"  
/db\_xref="GI:18873667"  
/translation="MEKVDMMTSGEGLQFSEKYQVYLSLAYSIIFILGPLNGTV  
LWHSWGQTKRWSCATTVLVNLMVADLVLLPFLITYSLDDRWPFGLKLLKLVHFLF  
YINLYSGTILLTCTISVHQFVCHPLCPSLYTRRHMLGTSTTWALVVLQPLTFLF  
SHTDYINGOMIYDWTSGENFDRLFAYGIVLTLSGFLSLGHFVFLFDGQEPDQARG  
EPHEDRQSPQVHPDHPGTGVWPLPLFCALPFLSHLLPFLHLLSAFSLPALDGSQQC  
LQDMEASGE"  
BASE COUNT 818 a 771 c 726 g 828 t  
ORIGIN  
Query Match 100.0%; Score 1082; DB 6; Length 3143;  
Best Local Similarity 100.0%; Pred. No. 1.7e-268;  
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCTGTCCATTTCCTTCCTCCAGGGAAGCAGAGCGGAGCGCTCGTGAGCTCTG 60  
Db 436 ATGCTGTCCATTTCCTTCCTCCAGGGAAGCAGAGCGGAGCGCTCGTGAGCTCTG 495  
Qy 61 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACAA 120  
Db 496 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACAA 555  
Qy 121 GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180  
Db 556 GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 615  
Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCCAAACC 240  
Db 616 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCCAAACC 675  
Qy 241 AAGCCTGGAGCTGTGCCACCACTATCTGTGGAACCTGTAGTGGTGGCCGACCTGCTTTAT 300  
Db 676 AAGCCTGGAGCTGTGCCACCACTATCTGTGGAACCTGTAGTGGTGGCCGACCTGCTTTAT 735  
Qy 301 GTGCTATTGGCCCTCTCTCATCATCACTACCTACTACCTAGATGACAGTGGCCCTTCGGGGAG 360  
Db 736 GTGCTATTGGCCCTCTCTCATCATCACTACTACCTAGATGACAGTGGCCCTTCGGGGAG 795  
Qy 361 GTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTACGGCAGCATCTCGCTG 420  
Db 796 GTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTACGGCAGCATCTCGCTG 855  
Qy 421 CTGACCTGCATCTCTGTGCACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 480  
Db 856 CTGACCTGCATCTCTGTGCACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 915  
Qy 481 TACCGGACCCGAGCAGTGCCTGGCTGGGCAACAGCACACCTGGGCCCTGTGTGCTCTC 540  
Db 916 TACCGGACCCGAGCAGTGCCTGGCTGGGCAACAGCACACCTGGGCCCTGTGTGCTCTC 975



	Qy	421	CTGACCTGCATCTCTGTGCAACCAGTTCCTTAGGTGTGTGCCACCCACTGTGGTCTCGCTGCC	480
	Db	123308	CTGACCTGCATCTCTGTGCAACCAGTTCCTTAGGTGTGTGCCACCCACTGTGGTCTCGCTGCC	123249
	Qy	481	TACCGGACCCGCAAGCATGCTTGCTGGGACACGAGCACACCTGGGCCCTGTGGTGTCTC	540
	Db	123248	TACCGGACCCGCAAGCATGCTTGCTGGGACACGAGCACACCTGGGCCCTGTGGTGTCTC	123189
	Qy	541	CAGCTGCTGCCACACCTGGGCTTTCTCCCACACGAGCTACATCAATGGCCAGATGATCTGG	600
	Db	123188	CAGCTGCTGCCACACCTGGGCTTTCTCCCACACGAGCTACATCAATGGCCAGATGATCTGG	123129
	Qy	601	TATGACATGACCAAGGAAGAATTGTGATCGGCTTTTTGCTTACGGCATAGTCTTGACA	660
	Db	123128	TATGACATGACCAAGGAAGAATTGTGATCGGCTTTTTGCTTACGGCATAGTCTTGACA	123069
	Qy	661	TTGFTCTGGCTTTCTTCCCTCCTTGGTCATTTTGGTGTGCTATTCTACTCATGGTCAGGAG	720
	Db	123068	TTGFTCTGGCTTTCTTCCCTCCTTGGTCATTTTGGTGTGCTATTCTACTCATGGTCAGGAG	123009
	Qy	721	CCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT	780
	Db	123008	CCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT	122949
	Qy	781	CCGGACCATCTACTGTGTGTGGCCCTTTACCCCTCTGTTTGTGGCCCTTCCATATCAC	840
	Db	122948	CCGGACCATCTACTGTGTGTGGCCCTTTACCCCTCTGTTTGTGGCCCTTCCATATCAC	122899
	Qy	841	TCGCTCCTTCTACCTCACATCTGCTTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGCG	900
	Db	122888	TCGCTCCTTCTACCTCACATCTGCTTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGCG	122829
	Qy	901	AGCCAGTGTGCGCTACAAGATATGGAGGCGCTCTGGTCACTGTGTGAGCAGCTGCTCAACC	960
	Db	122828	AGCCAGTGTGCGCTACAAGATATGGAGGCGCTCTGGTCACTGTGTGAGCAGCTGCTCAACC	122769
	Qy	961	AGTCCTCTACTTTCTTCAAGGGGGGCAAAAATAGAGTCAAGGCTCTCCAGAACTGAGG	1020
	Db	122768	AGTCCTCTACTTTCTTCAAGGGGGGCAAAAATAGAGTCAAGGCTCTCCAGAACTGAGG	122709
	Qy	1021	CAGAACAAAGTTGGTGAGCATCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT	1080
	Db	122708	CAGAACAAAGTTGGTGAGCATCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT	122649
	Qy	1081	GG 1082	
	Db	122648	GG 122647	
<hr/>				
	RESULT	5		PRI 30-JUL-2001
	LOCUS	AC092999	167084 bp	linear
	DEFINITION	Homo sapiens 3 BAC RP11-170K4 (Roswell Park Cancer Institute Human BAC library) complete sequence.		
	ACCESSION	AC092999		;
	VERSION	AC092999.3	GI:19033407	
	KEYWORDS	HTG.		
	SOURCE	human.		
	ORGANISM	Homo sapiens		
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 167084)		
	AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhav,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Bathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,		

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loiziged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.D., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Willamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (01-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (29-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Mar 1, 2002 this sequence version replaced gi:18139250.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,

unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	Source	Location/Qualifiers
misc_feature		1..167084 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-170K4" complement(1..2005) /note="overlaps bases 1..2005 of clone AC092997"
repeat_region		189..376 /rpt_family="MIR" complement(997..1123)
repeat_region		1999..2058 /rpt_family="L2"
repeat_region		2096..2404 /rpt_family="AluSx"
repeat_region		2422..2529 /rpt_family="L2"
repeat_region		2975..3004 /rpt_family="AT-rich"
repeat_region		3134..3163 /rpt_family="AT-rich"
repeat_region		4270..4290 /rpt_family="AT-rich"
repeat_region		complement(4421..4550) /rpt_family="MIR"
repeat_region		5251..5642 /rpt_family="MLT1K"
repeat_region		complement(5812..6366) /rpt_family="MER34B"
repeat_region		6909..6941 /rpt_family="AT-rich"
repeat_region		7595..7717 /rpt_family="MER5A"
repeat_region		complement(7788..8198) /rpt_family="MLT1I"
repeat_region		complement(8536..8591) /rpt_family="L2"
repeat_region		9634..9936 /rpt_family="AluSx"
repeat_region		complement(10766..10814) /rpt_family="MIR"
repeat_region		complement(10848..10964) /rpt_family="MIR"
repeat_region		17399..17431 /rpt_family="(TTAAA)n"
repeat_region		complement(17432..17546) /rpt_family="MIR"
repeat_region		17637..18073

Qy	541	CAGCTGTCGCCACACTGGCCTTCTCCACACAGCACTACATCAATGGCCAGATGATCTGG	600
Db	161461	CAGCTGTCGCCACACTGGCCTTCTCCACACAGCACTACATCAATGGCCAGATGATCTGG	161520
Qy	601	TATCAGATGACCCAGCCAAAGAGAATTTTGATGGCTTTTGGCTACGGCATAGTTCTGACA	660
Db	161521	TATCAGATGACCCAGCCAAAGAGAATTTTGATGGCTTTTGGCTACGGCATAGTTCTGACA	161580
Qy	661	TTGTCTGGCTTTCTTTTCCCTCCCTGGTGCATTTTGGTGTGCTATTCTACTGATGGTCAGGAG	720
Db	161581	TTGTCTGGCTTTCTTTTCCCTCCCTGGTGCATTTTGGTGTGCTATTCTACTGATGGTCAGGAG	161640
Qy	721	CCTGATCAAGCCAGAGAGAAACCTCATGAGGACAGGCACACAGCCGAGCCAGGTCAT	780
Db	161641	CCTGATCAAGCCAGAGAGAAACCTCATGAGGACAGGCACACAGCCGAGCCAGGTCAT	161700
Qy	781	CCGGACCATCCTACTGTGTGTGGCCCTTTCACCCCTCTGTTTTGTGCCCTTCCATATCAC	840
Db	161701	CCGGACCATCCTACTGTGTGTGGCCCTTTCACCCCTCTGTTTTGTGCCCTTCCATATCAC	161760
Qy	841	TCGCTCCTTCTACCTCACCATCTGCCTTTTCGCTTCTCAGGACTGCCAGCTCTTGATGGC	900
Db	161761	TCGCTCCTTCTACCTCACCATCTGCCTTTTCGCTTCTCAGGACTGCCAGCTCTTGATGGC	161820
Qy	901	AGCCAGTGTGGCCTTACAAAGATATGGAGGCCCTCTGGTGAGTGTGACAGCTGCCTCAACCC	960
Db	161821	AGCCAGTGTGGCCTTACAAAGATATGGAGGCCCTCTGGTGAGTGTGACAGCTGCCTCAACCC	161880
Qy	961	AGTCCTGTACTTTCTTTCAAGGGGGGCAAAAATAGAGTCAGGCTCCTCAGAAACTTGAGG	1020
Db	161881	AGTCCTGTACTTTCTTTCAAGGGGGGCAAAAATAGAGTCAGGCTCCTCAGAAACTTGAGG	161940

[illegible]



Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 10, 2000 this sequence version replaced gi:7248967.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
-----  
Center project name: L3302  
Center clone name: 12\_N\_17  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152766 bases at least Q40  
Consensus quality: 158075 bases at least Q30  
Consensus quality: 160495 bases at least Q20  
Insert size: 174000; agarose-fp  
Insert size: 162558; sum-of-contigs  
Quality coverage: 4.5 in Q20 bases; agarose-fp  
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 7254: contig of 7254 bp in length  
7255 7354: gap of 100 bp  
7355 8612: contig of 1258 bp in length  
8613 8712: gap of 100 bp  
8713 11463: contig of 2751 bp in length  
11464 11563: gap of 100 bp  
11564 14507: contig of 2944 bp in length  
14508 14607: gap of 100 bp  
14608 22452: contig of 7845 bp in length  
22453 22552: gap of 100 bp  
22553 35086: contig of 12534 bp in length  
35087 35186: gap of 100 bp  
35187 44278: contig of 9092 bp in length  
44279 44378: gap of 100 bp  
44379 56786: contig of 12408 bp in length  
56787 56886: gap of 100 bp  
56887 65270: contig of 8384 bp in length  
65271 65370: gap of 100 bp  
65371 76394: contig of 11024 bp in length  
76395 76494: gap of 100 bp  
76495 100320: contig of 23826 bp in length  
100321 100420: gap of 100 bp  
100421 113127: contig of 12707 bp in length  
113128 113227: gap of 100 bp  
113228 126756: contig of 13529 bp in length  
126757 126856: gap of 100 bp  
126857 151061: contig of 24205 bp in length  
151062 151161: gap of 100 bp  
151162 163958: contig of 12797 bp in length.

Location/Qualifiers  
1. .163958  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-12N17"  
/clone\_lbb="RPC1-11 Human Male BAC"

FEATURES  
source

misc\_feature 1. 7254  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
7355. .8612  
misc\_feature  
/note="assembly\_fragment"  
8713. .11463  
misc\_feature  
/note="assembly\_fragment"  
11564. .14507  
misc\_feature  
/note="assembly\_fragment"  
14608. .22452  
misc\_feature  
/note="assembly\_fragment"  
22553. .35086  
misc\_feature  
/note="assembly\_fragment"  
35187. .44278  
misc\_feature  
/note="assembly\_fragment"  
44379. .56786  
misc\_feature  
/note="assembly\_fragment"  
56887. .65270  
misc\_feature  
/note="assembly\_fragment"  
65371. .76394  
misc\_feature  
/note="assembly\_fragment"  
76495. .100320  
misc\_feature  
/note="assembly\_fragment"  
100421. .113127  
misc\_feature  
/note="assembly\_fragment"  
113228. .126756  
misc\_feature  
/note="assembly\_fragment"  
126857. .151061  
misc\_feature  
/note="assembly\_fragment"  
151162. .163958  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
BASE COUNT 49230 a 32369 c 32422 g 48536 t 1401 others  
ORIGIN  
Query Match 99.0%; Score 1071; DB 2: Length 163958;  
Best Local Similarity 99.9%; Pred. No. 1.6e-265;  
Matches 1082; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 ATGCTGTCATTTTTCCTTCCTCCAGGGGAAGCAGAGCGGAGCGCTCGTGAGCTCTG 60  
Db 98103 ATGCTGTCATTTTTCCTTCCTCCAGGGGAAGCAGAGCGGAGCGCTCGTGAGCTCTG 98162  
Qy 61 CTCCTGGAGGAGGCTCCCGGGACATGGAGAAGTGGAGATGAATACATACAGAAACAA 120  
Db 98163 CTCCTGGAGGAGGCTCCCGGGACATGGAGAAGTGGAGATGAATACATACAGAAACAA 98222  
Qy 121 GGTCTCTGCCAGTTCTCAGAGAAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180  
Db 98223 GGTCTCTGCCAGTTCTCAGAGAAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 98282  
Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCCTGCTGTGTGGCACTCTCTGGGCGCAACAC 240  
Db 98283 ATCTTTATCTAGGCTGCCACTAAATGGCCTGCTGTGTGGCACTCTCTGGGCGCAACAC 98342  
Qy 241 AAGCGCTGAGCTGTGCCACCACCTATCTGTGAACCTGTAGTGGCGGACCTGCTTTAT 300  
Db 98343 AAGCGCTGAGCTGTGCCACCACCTATCTGTGAACCTGTAGTGGCGGACCTGCTTTAT 98402  
Qy 301 GTGTAATGCCCTTCTCCTATCATCACCTACTACCTAGATGACAGGTGGCCCTTCGGGGAG 360  
Db 98403 GTGTAATGCCCTTCTCCTATCATCACCTACTACCTAGATGACAGGTGGCCCTTCGGGGAG 98462  
Qy 361 CTGCTGCAAGCTGGTGCACCTTCTGTCTTATATATACCTTTACGGCAGCATCTGCTG 420  
Db 98463 CTGCTGCAAGCTGGTGCACCTTCTGTCTTATATATACCTTTACGGCAGCATCTGCTG 98522  
Qy 421 CTGACCTGCATCTCTGTGCACAGTTCTCTAGGTGTGTGCCACCCACTGTGTTCGCTGCC 480  
Db 98523 CTGACCTGCATCTCTGTGCACAGTTCTCTAGGTGTGTGCCACCCACTGTGTTCGCTGCC 98582



QY 481 TACCGACCCGAGGATGCTGGTGGGACACAGCACACCTGGGCCCTGGTGGTCTTC 540  
Db 98583 TACCGACCCGAGGATGCTGGTGGGACACAGCACACCTGGGCCCTGGTGGTCTTC 98642  
QY 541 CAGCTGCTGCCACACTGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTGG 600  
Db 98643 CAGCTGCTGCCACACTGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTGG 98702  
QY 601 TATGACATGACACAGGCAAGAAATTTGATCGGCTTTTGGCCCTACGSCATAGTTCTGACA 660  
Db 98703 TATGACATGACACAGGCAAGAAATTTGATCGGCTTTTGGCCCTACGSCATAGTTCTGACA 98762  
QY 661 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTCTATTTCACTGATGCTCAGGAG 720  
Db 98763 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTCTATTTCACTGATGCTCAGGAG 98822  
QY 721 CTTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCCAT 780  
Db 98823 CTTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCCAT 98882  
QY 781 CCGGACCATCCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGGCTTCCATATCAC 840  
Db 98883 CCGGACCATCCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGGCTTCCATATCAC 98942  
QY 841 TCCTCCTTCTACCTCACCATCTGCTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGATGGC 900  
Db 98943 TCCTCCTTCTACCTCACCATCTGCTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGATGGC 99002  
QY 901 AG-CCAGTGTGGCTACAGATATGAGGCTCTGTTGAGTGTGAGCAGCTGCTCAACC 959  
Db 99003 AGCCAGTGTGGCTACAGATATGAGGCTCTGTTGAGTGTGAGCAGCTGCTCAACC 99062  
QY 960 CAGTCTGTACTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCCTCCAGAAACTGAG 1019  
Db 99063 CAGTCTGTACTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCCTCCAGAAACTGAG 99122  
QY 1020 GCAGACAAAGTGGGTGAGATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATC 1079  
Db 99123 GCAGACAAAGTGGGTGAGATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATC 99182  
QY 1080 TGG 1082  
Db 99183 TGG 99185

RESULT 7  
AF411108  
LOCUS AF411108 1019 bp DNA linear PRI 01-NOV-2001  
DEFINITION Homo sapiens GPR79 pseudogene, complete sequence.  
ACCESSION AF411108  
VERSION AF411108.1 GI:16566321  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1019)  
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,  
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.  
Discovery and mapping of ten novel G protein-coupled receptor genes  
Gene 275 (1), 83-91 (2001).  
JOURNAL MEDLINE 11574155  
PUBMED 21458557  
REFERENCE 2 (bases 1 to 1019)  
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,  
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.  
Direct Submission  
Submitted (17-AUG-2001) Department of Pharmacology, University of  
Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada  
JOURNAL  
FEATURES  
source 1. .1019  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

gene 1. .1019  
/gene="GPR79"  
/pseudo  
BASE COUNT 200 a 308 c 252 g 259 t  
ORIGIN  
Query Match 93.1%; Score 1007; DB 9; Length 1019;  
Best Local Similarity 99.9%; Pred. No. 3.6e-249;  
Matches 1018; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 ATGCTGCTCATTTTGTCTTCTCCAGGGGAAGCAGAGGGAGCGTGTGTGAGCTCTG 60  
Db 1 ATGCTGCTCATTTTGTCTTCTCCAGGGGAAGCAGAGGGAGCGTGTGTGAGCTCTG 60  
QY 61 CTCCTGAGGGAGCCCTCCGGGACATGGAGAGGTGGACATGAATACATCACAGGAACA 120  
Db 61 CTCCTGAGGGAGCCCTCCGGGACATGGAGAGGTGGACATGAATACATCACAGGAACA 120  
QY 121 GGTCTCTGCCAGTTCTCAGAGAGTAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGATATC 180  
Db 121 GGTCTCTGCCAGTTCTCAGAGAGTAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGATATC 180  
QY 181 ATCTTTATCCTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240  
Db 181 ATCTTTATCCTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240  
QY 241 AAGCGCTGGAGCTGTGCCACCACCTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300  
Db 241 AAGCGCTGGAGCTGTGCCACCACCTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300  
QY 301 GTGCTATTGCGCCCTTCCTCATCATCACTACTAGATGACAGGTGGCCCTTCGGGGAG 360  
Db 301 GTGCTATTGCGCCCTTCCTCATCATCACTACTAGATGACAGGTGGCCCTTCGGGGAG 360  
QY 361 CTGCTCTGAAGCTGTGACACTTCCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 420  
Db 361 CTGCTCTGAAGCTGTGACACTTCCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 420  
QY 421 CTGACCTGATCTCTGTGACACAGTTTCTAGGTGTGTGCCACCCACTGTCTTCGCTGCC 480  
Db 421 CTGACCTGATCTCTGTGACACAGTTTCTAGGTGTGTGCCACCCACTGTCTTCGCTGCC 480  
QY 481 TACCGACCCGAGGATGCTGGTGGGACACAGCACACCTGGGCCCTGGTGGTCTTC 540  
Db 481 TACCGACCCGAGGATGCTGGTGGGACACAGCACACCTGGGCCCTGGTGGTCTTC 540  
QY 541 CAGCTGCTGCCACACTGGCCCTTCTCCACACGACTACATCAATGGCCAGATGATCTGG 600  
Db 541 CAGCTGCTGCCACACTGGCCCTTCTCCACACGACTACATCAATGGCCAGATGATCTGG 600  
QY 601 TATGACATGACACAGGCAAGAAATTTGATCGGCTTTTGGCCCTACGSCATAGTTCTGACA 660  
Db 601 TATGACATGACACAGGCAAGAAATTTGATCGGCTTTTGGCCCTACGSCATAGTTCTGACA 660  
QY 661 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTCTATTTCACTGATGCTCAGGAG 720  
Db 661 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTCTATTTCACTGATGCTCAGGAG 720  
QY 721 CTTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCCAT 780  
Db 721 CTTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCCAT 780  
QY 781 CCGGACCATCCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGGCTTCCATATCAC 840  
Db 781 CCGGACCATCCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGGCTTCCATATCAC 840  
QY 841 TCCTCCTTCTACCTCACCATCTGCTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGATGGC 900  
Db 841 TCCTCCTTCTACCTCACCATCTGCTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGATGGC 900  
QY 901 AG-CCAGTGTGGCTACAGATATGAGGCTCTGTTGAGTGTGAGCAGCTGCTCAACC 959  
Db 901 AGCCAGTGTGGCTACAGATATGAGGCTCTGTTGAGTGTGAGCAGCTGCTCAACC 960

```
Qy 960 CAGTCTGTACTTTCTTCAAGGGGGCAAAATAGAGTCAAGGCTCCTCCAGAAACTGA 1018
|||||
Db 961 CAGTCTGTACTTTCTTCAAGGGGGCAAAATAGAGTCAAGGCTCCTCCAGAAACTGA 1019

RESULT 8
AX458308
LOCUS AX458308 1020 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 3 from Patent EP1215214.
ACCESSION AX458308
VERSION AX458308.1 GI:21725002
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Fidock, M.D.
TITLE Novel polypeptide
JOURNAL Patent: EP 1215214-A 3 19-JUN-2002;
Pfizer Limited (GB); PFIZER INC. (US)
FEATURES
Location/Qualifiers
source 1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 200 a 309 c 252 g 259 t
ORIGIN
Query Match 92.7%; Score 1002.8; DB 6; Length 1020;
Best Local Similarity 99.6%; Pred. No. 4.4e-248;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGCGGAGCGCTGTGGAGCTCTG 60
Db 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGCGGAGCGCTGTGGAGCTCTG 60

Qy 61 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 120
Db 61 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 120

Qy 121 GCTCTGCGCAATCTCAGAGAATCACAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 180
Db 121 GCTCTGCGCAATCTCAGAGAATCACAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 180

Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAAAAC 240
Db 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAAAAC 240

Qy 241 AAGCGCTGGAGCTGGCAACCACTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 300
Db 241 AAGCGCTGGAGCTGGCAACCACTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 300

Qy 301 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGGAG 360
Db 301 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGGAG 360

Qy 361 CTGCTCTGCAAGCTGGTGCACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 420
Db 361 CTGCTCTGCAAGCTGGTGCACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 420

Qy 421 CTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCCACCACTGTGTTCGCTGCC 480
Db 421 CTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCCACCACTGTGTTCGCTGCC 480

Qy 481 TACCGGACCCGACGATGCTGTGGTGGCCACGACCACTGGCGCTGGTGGTCTCTC 540
Db 481 TACCGGACCCGACGATGCTGTGGTGGCCACGACCACTGGCGCTGGTGGTCTCTC 540

Qy 541 CAGCTGTGCCACACGAGGCTTCTCCACACGAGCTACATCAATGCCAGATGATCTGG 600
Db 541 CAGCTGTGCCACACGAGGCTTCTCCACACGAGCTACATCAATGCCAGATGATCTGG 600
```

```
Qy 601 TATGACATGACACCAAGAGAAATTTTGATCGGCTTTTTCCTACGGCATAGTCTTGACA 660
|||||
Db 601 TATGACATGACACCAAGAGAAATTTTGATCGGCTTTTTCCTACGGCATAGTCTTGACA 660

Qy 661 TTGTCTGGCTTTCTTT--CCCTCCTTGTCTAFTTTTGGTGTCTATTCTACTGATGGTCAGG 718
|||||
Db 661 TTGTCTGGCTTTCTTTCCCTCCTTGTCTAFTTTTGGTGTCTATTCTACTGATGGTCAGG 720

Qy 719 AGCTGTATCAAGCCAGAGGAAACCTCATGAGGACAGCAACACAGCCCGAGCCAGGTCC 778
|||||
Db 721 AGCTGTATCAAGCCAGAGGAAACCTCATGAGGACAGCAACACAGCCCGAGCCAGGTCC 780

Qy 779 ATCCGACCATCTACTACTGTGGCTCTTCACCCCTCTGTTTGTGCCCTTCCATATC 838
|||||
Db 781 ATCCGACCATCTACTACTGTGGCTCTTCACCCCTCTGTTTGTGCCCTTCCATATC 840

Qy 839 ACTCGCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 898
|||||
Db 841 ACTCGCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 900

Qy 899 GCAGCAGTGTGGCTACAAGATATGAGGCGCTCTGGTGAAGTGTGAGCAGTGCCTCAAC 958
|||||
Db 901 GCAGCAGTGTGGCTACAAGATATGAGGCGCTCTGGTGAAGTGTGAGCAGTGCCTCAAC 960

Qy 959 CCAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGGCTCTCCAGAAACTGA 1018
|||||
Db 961 CCAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGGCTCTCCAGAAACTGA 1020

RESULT 9
AX375230
LOCUS AX375230 1076 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 20 from Patent WO0210387.
ACCESSION AX375230
VERSION AX375230.1 GI:19169948
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
Elliot, V.S., Rankumar, J., Baughn, M.R., Kallick, D.A., Walla, N.K.,
Hafalia, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Pollicky, J.L.,
Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0210387-A 20 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474806CB1"
BASE COUNT 211 a 318 c 274 g 273 t
ORIGIN
Query Match 92.7%; Score 1002.8; DB 6; Length 1076;
Best Local Similarity 99.6%; Pred. No. 4.4e-248;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTG 60
Db 57 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTG 116

Qy 61 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 120
Db 117 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 176

Qy 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTTACAGTATC 180
Db 177 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTTACAGTATC 236

Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAAAAC 240
```

|||||  
Db 237 ATCTTTATCTAGGGTGGCCACTAAATGGCACTGTCTTGTGGCACTCCCTGGGGCCAAACC 296  
QY 241 AAGCGCTGAGCTGTGCCACCACTATCTGTGTAACCTCATGTGTGGCCGACCTGCTTTAT 300  
Db 297 AAGCGTGGAGCTGTGCCACCACTATCTGTGTAACCTCATGTGTGGCCGACCTGCTTTAT 356  
QY 301 GTGCTATTGCCCTTCCCTATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360  
Db 357 GTGCTATTGCCCTTCCCTATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 416  
QY 361 GTGCTCTGAAGTGTGTGACCTTCTGTCTATATCAACCTTTACGGCAGCATCTGTGCTG 420  
Db 417 CTGCTCTGAAGTGTGTGACCTTCTGTCTATATCAACCTTTACGGCAGCATCTGTGCTG 476  
421 CTGACCTGATCTCTGTGCACCACTTCTAGTGTGTGCGCACCCACTGTGTGCTGCC 480  
Db 477 CTGACCTGATCTCTGTGCACCACTTCTAGTGTGTGCGCACCCACTGTGTGCTGCC 536  
QY 481 TACCGGACCCGAGGATGCTGTGGTGGCACCAGCACCACTGGGCCCTGTGTGCTC 540  
Db 537 TACCGGACCCGAGGATGCTGTGGTGGCACCAGCACCACTGGGCCCTGTGTGCTC 596  
QY 541 CAGCTGCTGCCACACTGGCTTCTCCACAGGACTATCAATGGCCAGATGATCTG 600  
Db 597 CAGCTGCTGCCACACTGGCTTCTCCACAGGACTATCAATGGCCAGATGATCTG 656  
QY 601 TATGACATGACCAAGCAAGAAATTTGATCGGCTTTTGGCTTATCTACTGATGTCGACA 660  
Db 657 TATGACATGACCAAGCAAGAAATTTGATCGGCTTTTGGCTTATCTACTGATGTCGACA 716  
QY 661 TTGCTGCGTCTTCTTT - CCCTCCTTGGTCAATTTGGTGTGCTATTTACTGATGTCAGG 718  
Db 717 TTGCTGCGTCTTCTTTCCCTTCTTGGTCAATTTGGTGTGCTATTTACTGATGTCAGG 776  
QY 719 AGCTGATCAAGCCAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTC 778  
Db 777 AGCTGATCAAGCCAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTC 836  
QY 779 ATCCGGACCATCTACTGTGTGTGGCTTCTCACCTCTGTGTGTGGCTTCCCATATC 838  
Db 837 ATCCGGACCATCTACTGTGTGTGGCTTCTCACCTCTGTGTGTGGCTTCCCATATC 896  
QY 839 ACTGCTCTTCTACTACCACTGCTCTGCTTCTGCTTCTCAGGACTGCCAGCTTCTGATG 898  
Db 897 ACTGCTCTTCTACTACCACTGCTCTGCTTCTGCTTCTCAGGACTGCCAGCTTCTGATG 956  
QY 899 GCAGCCAGTGTGCCCTACAGATATGAGGCTTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 958  
Db 957 GCACCCAGTGTGCCCTACAGATATGAGGCTTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 1016  
QY 959 CCAGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAAGGCTTCTCCAGAAACTGA 1018  
Db 1017 CCAGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAAGGCTTCTCCAGAAACTGA 1076

RESULT 10  
AX365514  
LOCUS AX365514 850 bp DNA linear PAT 15-FEB-2002  
DEFINITION Sequence 4 from Patent WO0185764.  
ACCESSION AX365514  
VERSION AX365514.1 GI:18697083  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Ramakrishnan,S.  
AUTHORS Regulation of human p2y-like g protein-coupled receptor  
TITLE Patent: WO 0185764-A 4 15-NOV-2001;  
JOURNAL Bayer Aktiengesellschaft (DE)  
FEATURES Location/Qualifiers

source 1..850  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 165 a 262 c 201 g 222 t  
ORIGIN  
Query Match 78.6%; Score 850; DB 6; Length 850;  
Best Local Similarity 100.0%; Pred. No. 1.3e-208;  
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 85 ATGGAGAAGGTGGACATGATACATCACAGGAACAAGGTCTCTGCCAGTCTCAGAGAAG 144  
Db 1 ATGGAGAAGGTGGACATGATACATCACAGGAACAAGGTCTCTGCCAGTCTCAGAGAAG 60  
QY 145 TACAACAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTTATCTTAGGGTGCACCTA 204  
Db 61 TACAACAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTTATCTTAGGGTGCACCTA 120  
QY 205 AATGGCACTGTCTTGTGGCACTCTCTGGGCGCAAAAGCGCTGGAGCTGTGCCACCA 264  
Db 121 AATGGCACTGTCTTGTGGCACTCTCTGGGCGCAAAAGCGCTGGAGCTGTGCCACCA 180  
QY 265 TATCTGGTGAACCTGATGTGGCGGACCTGCTTTATGTGCTATTGCCCTTCCCTCATCATC 324  
Db 181 TATCTGGTGAACCTGATGTGGCGGACCTGCTTTATGTGCTATTGCCCTTCCCTCATCATC 240  
QY 325 ACCTACTCACTAGATGACAGGTGGCCCTTTCGGGAGCTGCTCTGCAAGCTGTGTGCACTTC 384  
Db 241 ACCTACTCACTAGTACAGGTGGCCCTTTCGGGAGCTGCTCTGCAAGCTGTGTGCACTTC 300  
QY 385 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGTCATCTCTGTGCACTTC 444  
Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGTCATCTCTGTGCACTTC 360  
QY 445 TTCTAGTGTGTGGCCACCACTGCTCTGCTGCCCTACCGGACCGCGAGGATGCTCTGG 504  
Db 361 TTCTAGTGTGTGGCCACCACTGCTCTGCTGCCCTACCGGACCGCGAGGATGCTCTGG 420  
QY 505 CTGGGACCAAGCAACCACTGGGCTTCTGCTGCTCTCCAGCTGTGCTGCCACCACTGGCTTC 564  
Db 421 CTGGGACCAAGCAACCACTGGGCTTCTGCTGCTCTCCAGCTGTGCTGCCACCACTGGCTTC 480  
QY 565 TCCCAACAGGACTACATCAATAGGCGCAGATCTGCTGATGATGACATGACACCAAGAGAAT 624  
Db 481 TCCCAACAGGACTACATCAATAGGCGCAGATCTGCTGATGATGACATGACACCAAGAGAAT 540  
QY 625 TTTGATCGGCTTTTGCCTACGCGCATAGTCTGACATTTGCTGGCTTCTTCTCCCTCCTT 684  
Db 541 TTTGATCGGCTTTTGCCTACGCGCATAGTCTGACATTTGCTGGCTTCTTCTCCCTCCTT 600  
QY 685 GGTCATTTTGGTGTGCTATTCACTGATGGTCAGGAGCTGATCAAGCGAGAGAGAACCT 744  
Db 601 GGTCATTTTGGTGTGCTATTCACTGATGGTCAGGAGCTGATCAAGCGAGAGAGAACCT 660  
QY 745 CATGAGGACAGGCAACACAGCCCGAGCGAGCTCCATCCCGAGCACTCTACTGTGTGTGG 804  
Db 661 CATGAGGACAGGCAACACAGCCCGAGCGAGCTCCATCCCGAGCACTCTACTGTGTGTGG 720  
QY 805 CCTTCTCACCTCTGTTTGTGCTTCCATATCATCTGCTCTCTTACCTCACCACCTG 864  
Db 721 CCTTCTCACCTCTGTTTGTGCTTCCATATCATCTGCTCTCTTACCTCACCACCTG 780  
QY 865 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGAGCCAGCTGTGGCTTACAGATATG 924  
Db 781 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGAGCCAGCTGTGGCTTACAGATATG 840  
QY 925 GAGGCTCTG 934  
Db 841 GAGGCTCTG 850  
RESULT 11  
AC121579

**LOCUS** AC121579 209273 bp DNA linear HTG 20-MAY-2002  
**DEFINITION** Mus musculus chromosome UNK clone RP23-255H8, WORKING DRAFT  
**SEQUENCE**, 9 unordered pieces.  
**AC121579**  
**VERSION** AC121579.1 GI:20986651  
**KEYWORDS** HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
**SOURCE** Mus musculus  
**ORGANISM** Mus musculus  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
**AUTHORS** Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**TITLE** 1 (bases 1 to 209273)  
**JOURNAL** McPherson, J.D. and Waterston, R.H.  
**REFERENCE** The sequence of Mus musculus clone  
**UNPUBLISHED**  
**2** (bases 1 to 209273)  
**AUTHORS** McPherson, J.D. and Waterston, R.H.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
**COMMENT** Parkway, St. Louis, MO 63108, USA  
  

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
----- Project Information -----  
Center project name: M\_BA0255H8  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 205908 bases at least Q40  
Consensus quality: 206285 bases at least Q30  
Consensus quality: 206424 bases at least Q20  
Insert size: 177000; agarose-fp  
Insert size: 208741; sum-of-contigs  
Quality coverage: 15.45 in Q20 bases; agarose-fp  
Quality coverage: 12.20 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
  
1 1957: contig of 1957 bp in length  
\* 1958 2057: gap of unknown length  
\* 2058 7480: contig of 5423 bp in length  
\* 7481 7580: gap of unknown length  
\* 7581 21224: contig of 13644 bp in length  
\* 21225 21324: gap of unknown length  
\* 21325 40793: contig of 19469 bp in length  
\* 40794 40893: gap of unknown length  
\* 40894 73232: contig of 32339 bp in length  
\* 73233 73332: gap of unknown length  
\* 73333 113562: contig of 40230 bp in length  
\* 113563 113662: gap of unknown length  
\* 113663 163371: contig of 51709 bp in length  
\* 163372 165471: gap of unknown length  
\* 165472 208277: contig of 42806 bp in length  
\* 208278 209273: gap of unknown length  
\* 209274 209277: contig of 896 bp in length.  
  
Location/Qualifiers  
1. .209273  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP23-255H8"

**misc\_feature** 1. .1957  
**misc\_feature** /note="assembly\_name:Contig18"  
2058. .7480  
**misc\_feature** /note="assembly\_name:Contig19"  
7581. .21224  
**misc\_feature** /note="assembly\_name:Contig20"  
21325. .40793  
**misc\_feature** /note="assembly\_name:Contig21"  
40894. .73232  
**misc\_feature** /note="assembly\_name:Contig22"  
73333. .113562  
**misc\_feature** /note="assembly\_name:Contig23"  
113663. .165371  
**misc\_feature** /note="assembly\_name:Contig24"  
165472. .208277  
**misc\_feature** /note="assembly\_name:Contig25"  
208378. .209273  
**misc\_feature** /note="assembly\_name:Contig16"  
59147 a 42827 c 43921 g 62577 t 801 others  
  
BASE COUNT 59147 a 42827 c 43921 g 62577 t 801 others  
ORIGIN  
  
Query Match 60.0%; Score 649.4; DB 2; Length 209273;  
Best Local Similarity 78.3%; Pred. No. 1.3e-156;  
Matches 837; Conservative 0; Mismatches 211; Indels 21; Gaps 4;  
  
Qy 1 ATGCTGTCATTTTCTTCCAGGGAAGCAGAGGAGCGGCGCTGCTGGAGCTCTG 60  
Db 37890 ATGCTGTCGTTTATTTCTTTCAGGAGAACGAGAGTGCACGCC----- 37935  
  
Qy 61 CTCCTGGAGGAGCCTCCCGGACATGAGAGGTGGACATGATACATACACAGCAAA 120  
Db 37936 ATCTTCGAGGAAGCCTCTTTAGGATGGAGAGCTAGATATGAATGTCTCAAAGGTCGT 37995  
  
Qy 121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAAGACTACCTCTCCCTGGCCCTACAGTATC 180  
Db 37996 GCGCCTGCCACTTCTCAGAGAGTACAAAGACTACCTATCATTAACCTACAGTCTC 38055  
  
Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCCAAACC 240  
Db 38056 ATCTTTATCTTAGGATGCCCTCAATGGGCTGTCTTATGGCTCTCTGTGGTGCACAAACC 38115  
  
Qy 241 AAGCGCTGAGCTGTGCCACCACTATCTGTGACACCTGTGGTGGCGACCTGCTTTAT 300  
Db 38116 AAGCGCTGAGCTGTGCCACCACTATCTGTGACACCTGTGGTGGCGACCTGCTTTAT 38175  
  
Qy 301 GTGCTA---TTGCCCTTCTCATCATCACTACTACCTAGATGACAGGTGGCCCTTCGGG 357  
Db 38176 ATGCTAAGCCTGCCCTTCTCATCATCACTACCTACCGGTTGGTGACACCTGGCTCTTTGGG 38235  
  
Qy 358 GAGCTGCTCTCAAGCTGGTGCACCTTCTCTATATCAACCTTTACGGCAGCATCTG 417  
Db 38236 GAGATGCTGTGAGGTGGTGGCTTCTGTCTACACCAACCTCTACAGCAGCATCTG 38295  
  
Qy 418 CTGCTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCACCCTCTGTGTTCGCTG 477  
Db 38296 CTGCTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCACCCTCTGTGTTCGCTG 38355  
  
Qy 478 CCTTACCGGACCGGAGCATGCTGGTGGGACACACACCTGGGCCCTGTGTGTC 537  
Db 38356 TCCTACCGGACTCACCGGACCGCTGGCAGCAGCTGTGTGGCACCCTGGTGTGTT 38415  
  
Qy 538 CTCAGCTGCTGCCACACTGGCTTCTCCACACGGGACTACATCAATGGCCAGATGATC 597  
Db 38416 CTCCAATTTGCTGGCAGCTCTGTGTCTTCTCCACACAGACTATGTAATGGCCAGGATGATC 38475  
  
Qy 598 TGGTATGACATGACCAAGAGAAATTTGATCGGCTTTTGTGCTAGCGGATAGTCTG 657  
Db 38476 TGGTATGACATGACCAAGAGAAATTTGATCGGCTTTTGTGCTAGCGGATAGTCTG 38535  
  
Qy 658 ACATTGCTGGCTT--TCTTTCCCTCCTTGGTCAATTTGGTGTGCTATCACTGATGTC 715  
Db 38536 ACATTGCTGGCTT--TCTTTCCCTCCTTGGTCAATTTGGTGTGCTATCACTGATGTC 38595

TITLE		REFERENCE	AUTHORS	TITLE	JOURNAL		REFERENCE	AUTHORS	TITLE	JOURNAL		COMMENT
-------	--	-----------	---------	-------	---------	--	-----------	---------	-------	---------	--	---------

```
* 18320 19563: contig of 1244 bp in length
* 19564 19663: gap of unknown length
* 19664 21323: contig of 1660 bp in length
* 21324 21423: gap of unknown length
* 21424 22609: contig of 1186 bp in length
* 22610 22709: gap of unknown length
* 22710 24220: contig of 1511 bp in length
* 24221 24320: gap of unknown length
* 24321 25466: contig of 1146 bp in length
* 25467 25566: gap of unknown length
* 25567 27297: contig of 1731 bp in length
* 27298 27397: gap of unknown length
* 27398 28793: contig of 1396 bp in length
* 28794 28893: gap of unknown length
* 28894 30289: contig of 1396 bp in length
* 30290 30389: gap of unknown length
* 30390 31748: contig of 1357 bp in length
* 31749 31848: gap of unknown length
* 31849 33603: contig of 1757 bp in length
* 33604 33703: gap of unknown length
* 33704 35261: contig of 1558 bp in length
* 35262 35361: gap of unknown length
* 35362 37315: contig of 1954 bp in length
* 37316 37415: gap of unknown length
* 37416 38787: contig of 1372 bp in length
* 38788 38887: gap of unknown length
* 38888 40559: contig of 1672 bp in length
* 40560 40659: gap of unknown length
* 40660 42655: contig of 1996 bp in length
* 42656 42755: gap of unknown length
* 42756 44163: contig of 1407 bp in length
* 44164 44263: gap of unknown length
* 44264 46500: contig of 2238 bp in length
* 46501 46600: gap of unknown length
* 46601 48572: contig of 1972 bp in length
* 48573 48672: gap of unknown length
* 48673 50544: contig of 1872 bp in length
* 50545 50644: gap of unknown length
* 50645 51883: contig of 1239 bp in length
* 51884 51983: gap of unknown length
* 51984 54568: contig of 2585 bp in length
* 54569 54668: gap of unknown length
* 54669 56243: contig of 1577 bp in length
* 56246 56345: gap of unknown length
* 56346 58742: contig of 2397 bp in length
* 58743 58842: gap of unknown length
* 58843 61978: contig of 3136 bp in length
* 61979 62078: gap of unknown length
* 62079 63933: contig of 1855 bp in length
* 63934 64033: gap of unknown length
* 64034 65836: contig of 1803 bp in length
* 65837 65936: gap of unknown length
* 65937 68259: contig of 2323 bp in length
* 68260 68359: gap of unknown length
* 68360 70942: contig of 2583 bp in length
* 70943 71042: gap of unknown length
* 71043 74577: contig of 3535 bp in length
* 74578 74677: gap of unknown length
* 74678 76364: contig of 1687 bp in length
* 76365 76464: gap of unknown length
* 76465 80009: contig of 3545 bp in length
* 80010 80109: gap of unknown length
* 80110 85962: contig of 5853 bp in length
* 85963 86062: gap of unknown length
* 86063 91107: contig of 5045 bp in length.
FEATURES             Location/Qualifiers
     source            1..91107
                        /organism="Rattus norvegicus"
                        /db_xref="taxon:10116"
                        /clone="CH230-18307"
BASE COUNT      25705 a 18632 c 18796 g 23400 t 4574 others
ORIGIN
```

```
Query Match          59.6%; Score 644.4; DB 2; Length 91107;
Best Local Similarity 78.2%; Pred. No. 2.4e-155;
Matches 832; Conservative 0; Mismatches 211; Indels 21; Gaps 4;

Qy 1 ATGCTGCCATTTTTCCTTCAGGGGGAAGCAGGAGCGGAGCCGTCGTGGAGCTCTG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78493 ATGCTGTGCGTTTATTTCCTTACAGGAGCAGCAGAACTAGAGGCC----- 78448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 CTCCTGGAGGAGCCTCCCGGACATGGAAGGTGGACATGAATACATACAGCAACA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78447 ATCTTTGAGGAAGCCTCGTTGAGATGGAGAGCTAGACGTGAATGCTCAGAGGATGAA 78388
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 GGTCTCTGCCAGTTTCTCAGAGAACTACAAAGTCTACCTCTCCCTGGCCCTACAGTATC 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78387 GCCTTCGCCACTTCTCGGAGAGCTACAAAGCAAGTCTACCTATCGCTAACCTACAGTCTC 78328
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGCTTTGTGGCACTCTCTGGGCAAAACC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78327 ATCTTATGCTAGGAATGCCCTCAATGGCGCGCTTATGGCTCTCTCTGGTGCCAAACC 78268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 241 AAGCGCTGGAGCTGGCCACCACCTATCTGCTGAACCTGATGGTGGCGGACCTGCTTTAT 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78267 AAGCCCTGGAGCTGTACCACCACCTACTACCTGATGAACCTGATGGTGGCGGACCTGCTTAC 78208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 301 GTGCT---ATTGCCCTTCTCTCATCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGG 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78207 ATGCTGAGTCTGCCCTTCTCTCATCATCACTACGCGGGGGTGACACCTGCTCTTTGGG 78148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 358 GAGCTGCTCTCAAGCTGGTGCACCTTCTCTTATATCAACCTTTACGGCAGCATCTCG 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78147 GAGAGCTGTGACGCTGGTGGCTTCTCTTCTACCAACCTCTACAGCAGCATCTCG 78088
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 418 CTGCTGACCTGCATCTCTGTGCACACAGTCTCTAGTGTGTGGCCACCTACTGTGTGCTG 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78087 CTGCTGACCTGCATCTCTCTATACCGCTTCTAGAGGTGGCCACCTGAGCTCACTG 78028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 478 CCCTACCGGACCGGAGCATGCTGGCTGGCGACACACACCTGGGCGCCCTGGTGTC 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78027 TCCTTCCGGACCCACCGGCATACCTGGCAAGCAGCGGCACCTGGACCTGGTATG 77968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 538 CTCAGCTGTGCCACACTGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATC 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77967 GTCCAAATGTCTACCGACTCTGTCTTCTCCACACAGACTATGTAAATGGCCACCTGATC 77908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 598 TGGTATGACATGACGCCAAGAAATTTGATCGGCTTTTGGCTACGGCATAGTCTTG 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77907 TGCTACGACACAGCCAGCCAGAGAAATTCGATCATTTCTTCGCCCTACAGCATGGTGCTG 77848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 658 ACATTGTCTGGCTT--TCCTTTCCCTCTTGGCTCATTTTGGTGTGCTATTCTACTGATGCTC 715
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77847 ACATTATCTGGGTTTGTCTTCCCTTCCCTTGTATCATTTCTCGGCTGCTACTCTCTGATGTC 77788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 716 AGGAGCTGTATCAAGCCAGAGGAGAACTCATAGGACAGCAACACACCGCCAGCCAGG 775
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77787 AGGAGCTGTACAAATCCAGCAGAGAGCCCTCCCAAGCAGACGCCACCGCCGTCGAAG 77728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 776 TCCATCCGGACCATCTACTACTGTGTGGCGCTCTTTCACCCCTCTGTTTTGTGCCCTTCCAT 835
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77727 TCTATCCGGACCATCTCTGCTGTGTAGCCTCTTTCATGCTCTGTTTTGTGCCCTTCCAC 77668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 836 ATCACTGCTCTCTTCTACTCTACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTG 895
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77667 ATTGCCGGTCTCTTCTACTTGTGTCATCGCTTTCTCTGAACTCACAGGATGGCAGCTCTTG 77608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 896 ATGCAGCCAGTGTGGCTCAAGATATGAGGCGCTCTGTGTGAGTGTGAGCAGCTGCCTC 955
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77607 TTGCAGCCAGCATGGCTCAAGATCTGTAGGCTCTGTGTAGCGCAGCAGCTGCGCTC 77548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 956 AACCCAGTCTGTACTTTTCAAGGGGGCAAAAATAGAGTCAAGGCTCTCCAGAAAC 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77547 AACCCGGTCTGTACTTCTTTGTCA--CAGGGCAACAACAGAGTCCAACCTCTCCCAAGC 77490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1016 TGAGGCAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAAGA 1059
```



LOCUS	AC120742	192524 bp	DNA	linear	HTG 24-JUL-2002





```
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300
|||||
Db 159506 AAGCGCTGGAGCTGTACCACCATCTACCTGATGAACCTGATGTGGCCGACCTGCTTTAC 159447
QY 301 GTGCT---ATTGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGG 357
|||||
Db 159446 ATGCTGAGTCTGCCCTTCCTCATCATCACTACCTACCGGGGGTGACACCTTGGCTCTTTGGG 159387
QY 358 GAGCTGCTGTGAAGCTGGTGCACTTCCCTGTTCTATATCAACCTTTTACGGCAGCATCCTG 417
|||||
Db 159386 GAGACGCTGTGAGGCTGGTGCGCTTCTGTTCTTACACCAACCTCTACAGCAGCATCCTG 159327
QY 418 CTGCTGACCTGCATCTCTGTGCAACCACTTCTAGGTGTGTGCCACCACTGTGTTCGCTG 477
|||||
Db 159326 CTGCTGACCTGCATCTCTCTATACCGCTTCTTAGGAGTGTGCCACCACTGAGCTCACTG 159267
QY 478 CCTACCGGACCGCAGGAGCATGCTGGTGGGACACAGCAGCAGCTGGGCGCTTGGTGGTC 537
|||||
Db 159266 TCCCTCCGGACCGCAGGCACTAGCTGGCAAGCAGCGCCACCACTGGAGCTGGTAGTC 159207
QY 538 CTCACGCTGCTGCCACACCTGGCTTCTCCACACGCACTACATCAATGCCAGATGATC 597
|||||
Db 159206 GTCCAATTGCTACCGACTCTGTCTTCTCCACACAGACTATGTAATGGCCACGCTGATC 159147
QY 598 TGGTATGACATGACAGCCCAAGAGAAATTTTGATCGGCTTTTTCCTACGGCATAGTTCTG 657
|||||
Db 159146 TGCTAGCACACAGCCAGCCAGAGAAATTCGATCATTTCTTCGCTTACAGCATGGTCTG 159087
QY 658 ACATTTGCTGGCTT--TCCTTTCCTCTGCTGGTCAATTTTGGTGTGCTATTCACATGATGGTC 715
|||||
Db 159086 ACATATCTGGTGTGCTTTCCTCCCTCCTTGATCATTTCTCGGCTGCTACTCTCTGATGGTC 159027
QY 716 AGGAGCTGATCAAGCAGAGGAGAACCTCATGAGGACAGGCAACAGCCCGGAGCCAGG 775
|||||
Db 159026 AGGAGCTGACAATTCAGCAGAGGAGCTCCCAACGACAGGCGCCAGCGCCGCTGGCCAAG 158967
QY 776 TCCATCCGGACCATCTCTACTGGTG 799
|||||
Db 158966 TCTATCCGGACCATCTCTGCTGGTG 158943
```

```
RESULT 15
AX244718 681 bp DNA linear PAT 28-SEP-2001
Sequence 47 from Patent WO0166750.
ACCESSION AX244718
VERSION AX244718.1 GI:15859597
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Vogeli, G. and Wood, L.S.
TITLE G protein-coupled receptors
JOURNAL Patent: WO 0166750-A 47 13-SEP-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 163 a 171 c 171 g 176 t
ORIGIN

Query Match 36.6%; Score 396.4; DB 6; Length 681;
Best Local Similarity 99.7%; Pred. No. 1.8e-91;
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 685 GGTCAATTTTGGTGTGCTATTTCACCTGATGGTCAGGAGCCTGATCAAGCCAGGAGGAGACCT 744
|||||
Db 681 GGTCACTTTTGGTGTGCTATTTCACCTGATGGTCAGGAGCCTGATCAAGCCAGGAGGAGACCT 622
```

```
QY 745 CATGAGGACAGGCAACACAGCCCGAGCGAGGTCCATCCGACCATCTACTGTGTGTGG 804
|||||
Db 621 CATGAGGACAGGCAACACAGCCCGAGCGAGGTCCATCCGACCATCTACTGTGTGTGG 562
QY 805 CCTCTTACACCTCTGTTTGTGCCCCCTTCCATATCATCTACCTCGCTCCCTTCTACCTCAGCATCTG 864
|||||
Db 561 CCTCTTACACCTCTGTTTGTGCCCCCTTCCATATCATCTACCTCGCTCCCTTCTACCTCAGCATCTG 502
QY 865 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGGAGCCAGTGTGGCCTACAGATATG 924
|||||
Db 501 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGGAGCCAGTGTGGCCTACAGATATG 442
QY 925 GAGGCCCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTGTACTTTCTTTCAAGGGG 984
|||||
Db 441 GAGGCCCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTGTACTTTCTTTCAAGGGG 382
QY 985 GGCAAAAATAGATCAGGCTCCTCCAGAAACTGAGGCGAGAACAAAGTTGGTGAGCATCCA 1044
|||||
Db 381 GGCAAAAATAGATCAGGCTCCTCCAGAAACTGAGGCGAGAACAAAGTTGGTGAGCATCCA 322
QY 1045 GCTGGGAGGAGAGATGCCCGAGGCTTGAACAGATCTGG 1082
|||||
Db 321 GCTGGGAGGAGAGATGCCCGAGGCTTGAACAGATCTGG 284
```

Search completed: May 26, 2003, 22:35:37  
Job time : 3961 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 20:12:01 ; Search time 311 Seconds  
(without alignments)  
7834.927 Million cell updates/sec

Title: US-10-023-586B-1

Perfect score: 1082

Sequence: 1 atgtgttcatttcttctcc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1082	100.0	3143	24 AAS17746	Human genomic clone
2	1051	97.1	1063	24 AAS98135	Human DNA for pote
3	1002.8	92.7	1076	24 AAD29667	Human G-protein co
4	996.4	92.1	1002	22 AAH73516	Human G-protein-co
5	850	78.6	850	24 AAH73516	Human P2Y-like G p
6	510	47.1	510	22 AAH50975	Human nGPCR12 codi
7	396.4	36.6	681	22 AAS30774	Human cDNA encodin
8	370	34.2	585	22 AAS30783	Human cDNA encodin
9	227.8	21.1	2025	22 AAD04980	Human purinergic r

10	223.6	20.7	1163	22 AAD04981	Human purinergic r
11	222	20.5	551	22 AAS42837	Human G Protein-Co
12	202.2	18.7	1842	16 AAQ88134	Human P20 receptor
13	202.2	18.7	2138	24 AAD32337	Mus musculus GPCR
14	202.2	18.7	2138	24 AAD32944	Mus musculus GPCR
15	178.6	16.5	1429	18 AAT74321	Human P2Y4 recepto
16	170	15.7	234	22 AAS30782	Human cDNA encodin
17	164.8	15.2	984	17 AAT18368	Human placenta G-p
18	156.8	14.5	984	18 AAT70000	P2U purinergic rec
19	156.8	14.5	984	18 AAT47730	Human placenta pur
20	154	14.2	593	23 AAS66708	DNA encoding novel
21	148.8	13.8	984	17 AAT18367	Mouse pancreas G-p
22	105.8	9.8	6721	24 AAS18599	Purinergic recepto
23	105.4	9.7	6721	24 AAS18600	Purinergic recepto
24	97.8	9.0	1543	24 ABK12957	DNA sequence of mo
25	97.4	9.0	1119	23 ABT98009	Non-endogenous hum
26	95.6	8.8	1892	24 AAT72351	DEZ orphan recepto
27	92	8.5	1477	16 AAQ84557	Murine C140 recept
28	92	8.5	1477	17 AAT32036	Murine C140 recept
29	92	8.5	2732	16 AAQ84559	Murine C140 recept
30	92	8.5	2732	17 AAT32038	Murine C140 recept
31	91.4	8.4	1176	23 ABT98024	Non-endogenous hum
32	91.4	8.4	1634	14 AAQ45553	Human somatostatin
33	90.6	8.4	9409	22 AAQ76477	Human immune/haema
34	89	8.2	1526	24 ABK94922	Human novel polynu
35	88.8	8.2	1001	24 AAS98131	Human DNA for pote
36	88.4	8.2	450	21 AAA42673	Human secreted exp
37	88.4	8.2	1020	21 AAA30601	Human G protein-co
38	88.4	8.2	1020	21 AAA30718	DNA encoding human
39	88.4	8.2	1020	24 ABK52848	Human cysteinyl le
40	88.4	8.2	1901	15 AAQ66178	Seven transmembran
41	88.4	8.2	1901	19 AAV18356	Human R12 seven tr
42	88.4	8.2	1901	21 AAA91725	Human 7TM receptor
43	88.4	8.2	1901	24 ABK54254	Human 7 transmembr
44	88.4	8.2	2453	18 AAT44092	Human G-protein th
45	87.4	8.1	1119	21 AAA30626	Human G protein-co

ALIGNMENTS

RESULT 1

AAS17746	
ID AAS17746 standard; DNA; 3143 BP.	
XX	
AC AAS17746;	
XX	
DT 26-FEB-2002 (first entry)	
XX	Human genomic clone for P2Y-like G protein-coupled receptor.
DE	
DE	
XX	
KW	Human; ds: P2Y-like G protein-coupled receptor; GPCR; COPD;
KW	Chronic obstructive pulmonary disease; nervous system disease;
KW	Parkinson's disease; multiple sclerosis; dementia; stroke;
KW	Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;
KW	Bacterial infection; fungal infection; protozoan infection;
KW	viral infection; pain; cancer; anorexia; bulimia; asthma;
KW	acute heart failure; hypotension; hypertension; osteoporosis; diabetes;
KW	angina pectoris; myocardial infarction; ulcer; inflammation; allergy;
KW	psychotic disorder; neurological disorder; anxiety; schizophrenia;
KW	manic depression; delirium; severe mental retardation; dyskinesia.
XX	
OS Homo sapiens.	
XX	
FH	Key
FT	Location/Qualifiers
FT	520..2498
FT	/*tag= a
FT	/product= "P2Y-like GPCR"
XX	
PN	WO200185764-A2.
XX	
PD	15-NOV-2001.
XX	

PF 09-MAY-2001; 2001WO-EP05244.  
XX  
PR 11-MAY-2000; 2000US-203582P.  
XX 21-FEB-2001; 2001US-269857P.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Ramakrishnan S;  
XX  
DR WPI: 2002-075242/10.  
DR P-PSDB; AAU11451.  
XX  
PT New polynucleotides for producing P2Y-like G protein-coupled receptors  
PT (GPCR) that are used for screening inhibitors or regulators of human  
PT P2Y-like GPCR, especially useful for treating pain, cancer or  
PT neurological disorders  
XX  
PS Disclosure: Fig 1: 114pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a P2Y-like  
CC G protein-coupled receptor (GPCR) polypeptide, its fragment,  
CC derivative or allele, a host cell containing an expression vector  
CC comprising the polynucleotide and screening for agents that regulate the  
CC GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR  
CC polypeptide, which may be employed for screening agents that inhibit or  
CC regulate human P2Y-like GPCR. The reagent or inhibitor of the human  
CC P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR  
CC disorders, particularly COPD (chronic obstructive pulmonary disease),  
CC peripheral or central nervous system disease (e.g. Parkinson's  
CC disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and  
CC many other diseases and disorders listed in the specification), benign  
CC prostatic hyperplasia or urinary incontinence. A pharmaceutical  
CC composition containing the modulators and/or regulators of P2Y-like  
CC GPCR is useful for modulating the activity of a P2Y-like GPCR.  
CC In particular, these are useful for treating, preventing or ameliorating  
CC infections (e.g. bacterial, fungal, protozoan or viral infections), pain,  
CC cancer, anorexia, bulimia, asthma, acute heart failure, hypotension,  
CC hypertension, osteoporosis, diabetes, angina pectoris, myocardial  
CC infarction, ulcers, inflammation, allergies, psychotic or neurological  
CC disorders (e.g. anxiety, schizophrenia, manic depression, delirium,  
CC severe mental retardation or dyskinesias). The present sequence is  
CC a genomic clone encoding the P2Y-like GPCR of the invention.  
XX  
SQ Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 other;  
  
Query Match 100.0%; Score 1082; DB 24; Length 3143;  
Best Local Similarity 100.0%; Pred. No: 1.9e-284;  
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGCTGTCCATTTTGGTCTTCCAGGGGAGCAGAGCGGGAGCGTCTGGAGCTCTG 60  
DB 436 ATGCTGTCCATTTTGGTCTTCCAGGGGAGCAGAGCGGGAGCGTCTGGAGCTCTG 495  
  
61 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA 120  
DB 496 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA 555  
  
121 GGTCTCTGCCAGTCTTCAGAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 180  
DB 556 GGTCTCTGCCAGTCTTCAGAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 615  
  
181 ATCTTTATCTAGGGTGGCCACTTAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCC 240  
DB 616 ATCTTTATCTAGGGTGGCCACTTAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCC 675  
  
241 AAGCGCTGAGCTGTGCCACCACTATCTCTGTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300  
DB 676 AAGCGCTGAGCTGTGCCACCACTATCTCTGTGGTGAACCTGATGTGGCCGACCTGCTTTAT 735  
  
301 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360  
DB 736 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 795

QY 361 CTGCTCTGCAAGCTGTGTGCACTTCTGTCTTATATCAACCTTTACGCGCAGCATCTCTGCTG 420  
DB 796 CTGCTCTGCAAGCTGTGTGCACTTCTGTCTTATATCAACCTTTACGCGCAGCATCTCTGCTG 855  
  
QY 421 CTGACCTGATCTCTGTGTGCAAGTCTCTAGGTGTGTGTCACCCACTGTGTGCGTGCCTC 480  
DB 856 CTGACCTGATCTCTGTGTGCAAGTCTCTAGGTGTGTGTCACCCACTGTGTGCGTGCCTC 915  
  
QY 481 TACCGGACCCGAGGATGCTGTGGTGGCCACGACACCACTGGGCGCTGTGTGCTCTC 540  
DB 916 TACCGGACCCGAGGATGCTGTGGTGGCCACGACACCACTGGGCGCTGTGTGCTCTC 975  
  
QY 541 CAGCTGTCTGCCACACTGTGCTCTTCTCCACAGGACTACATCAATGCGCAGATGATCTGG 600  
DB 976 CAGCTGTCTGCCACACTGTGCTCTTCTCCACAGGACTACATCAATGCGCAGATGATCTGG 1035  
  
QY 601 TATGACATGACACGACCAAGAGAAATTTTGTGCGCTTTTTCCTACGGCATAGTCTTGACA 660  
DB 1036 TATGACATGACACGACCAAGAGAAATTTTGTGCGCTTTTTCCTACGGCATAGTCTTGACA 1095  
  
QY 661 TTGTCTGGCTTTCTTCCCTCTTGTGTCTTGTGTCTTATTTGCTACTGATGCTCAGGAG 720  
DB 1096 TTGTCTGGCTTTCTTCCCTCTTGTGTCTTGTGTCTTATTTGCTACTGATGCTCAGGAG 1155  
  
QY 721 COTGATCAAGCAGAGGAGAACTCATGAGGACAGCAACAGCCGAGCCAGGTCCAT 780  
DB 1156 COTGATCAAGCAGAGGAGAACTCATGAGGACAGCAACAGCCGAGCCAGGTCCAT 1215  
  
QY 781 CCGGACCACTCTACTGCTGTGTGGCTCTTCAACCTCTCTTTGTGCGCTTCCATATCAC 840  
DB 1216 CCGGACCACTCTACTGCTGTGTGGCTCTTCAACCTCTCTTTGTGCGCTTCCATATCAC 1275  
  
QY 841 TCGCTCTTCTTACCTCACCATCTGCTTTCTGCTTCTTCAGGACTGCCAGCTCTTGTATGGC 900  
DB 1276 TCGCTCTTCTTACCTCACCATCTGCTTTCTGCTTCTTCAGGACTGCCAGCTCTTGTATGGC 1335  
  
QY 901 AGCCAGTGTGGCTTCAAGATATGAGGCTCTGTGTGCTGTGAGTGAGAGCTGCTCAACCC 960  
DB 1336 AGCCAGTGTGGCTTCAAGATATGAGGCTCTGTGTGCTGTGAGTGAGAGCTGCTCAACCC 1395  
  
QY 961 ACTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG 1020  
DB 1396 ACTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG 1455  
  
QY 1021 CAGAACAAAGTGTGGTGAGCATCCAGCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 1080  
DB 1456 CAGAACAAAGTGTGGTGAGCATCCAGCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 1515  
  
QY 1081 GG 1082  
DB 1516 GG 1517  
  
RESULT 2  
AAS98135  
ID AAS98135 standard; DNA; 1063 BP.  
XX  
AC AAS98135;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Human DNA for potential G protein-coupled receptor #92.  
XX  
KW Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;  
KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;  
KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;  
KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;  
KW depression; epilepsy; macular degeneration; lymphoma; melanoma;  
KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;  
KW psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;  
KW tuberculosis; cognition disorder; memory disorder; anorexia;  
KW hormonal release disorder; cardiovascular activity disorder;  
KW pain perception disorder; obesity; diabetes; obesity;

diabetes; hyperlipidaemia; stroke; gene therapy.

Homo sapiens.

W0200185791-A1.

15-NOV-2001.

11-MAY-2001; 2001WO-US15332.

11-MAY-2000; 2000US-203217P.

18-MAY-2000; 2000US-205945P.

(LIFE-) LIFESPAN BIOSCIENCES INC.

Brown JP, Millier M, Burmer G, Fabre-Suver C, Pritchard D;

WPI; 2002-066595/09.

Novel G protein-coupled receptor polypeptides including galanin receptor polypeptides useful for identifying modulators that are useful for treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis, stroke

Claim 2; Page 98; 144pp; English.

The invention relates to an isolated polypeptide encoded by a nucleic acid molecule that is at least 80% identical to the G protein-coupled (GPCR) polynucleotides included in the specification. Also included are probes based on the GPCR sequences (including antisense probes), a host cell comprising an expression vector comprising the GPCR sequence, antibodies raised against the polypeptides, and methods of identifying modulators of the polypeptides. The polypeptides are useful for identifying modulator compounds which function as modulators, activators, repressors, agonists or antagonists of the novel GPCR polypeptides including the GAL4 polypeptide. The antibodies and nucleic acid probes as described above can be used to detect the presence of the polypeptides and nucleic acids and are used to diagnose a variety of diseases or disorders in which GPCRs are involved e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma, atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease, depression, epilepsy, macular degeneration, lymphoma, melanoma, multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis, tuberculosis and many other diseases listed in the specification. The probes and antibodies are also useful for diagnosing cognition and memory disorders, anorexia, hormonal release disorders, cardiovascular activity disorders, pain perception disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds that decrease or increase the expression of galanin receptor (GAL4) can be used to treat obesity, diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is useful for treating the above mentioned disorders by gene therapy techniques. The present sequence is a novel GPCR polynucleotide of the invention.

Sequence 1063 BP; 216 A; 315 C; 272 G; 260 T; 0 other;

Query Match 97.1%; Score: 1051; DB 24; Length 1063;  
Best Local Similarity 99.9%; Pred. No. 3.3e-276;  
Matches 1062; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 19 CCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTGCTCTCGGAGGAGCGCTCC 78  
DB 1 CCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTGCTCTCGGAGGAGCGCTCC 60  
QY 79 CGGACATGAGAGGTGGACATATACATACAGGAACAAGGTCTCTCCAGTTCTCA 138  
DB 61 CGGACATGAGAGGTGGACATATACATACAGGAACAAGGTCTCTCCAGTTCTCA 120  
QY 139 GAGAAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTTATCTAGGCTG 198  
DB 121 GAGAAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTTATCTAGGCTG 180

QY 199 CCACTAAATGGCACTGTCTTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCC 258  
DB 181 CCACTAAATGGCACTGTCTTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCC 240  
QY 259 ACCACCTATCTGTGTGAACCTGATGGTGGCGACCTGCTTTATGTGTATTTGCGCTTCCTC 318  
DB 241 ACCACCTATCTGTGTGAACCTGATGGTGGCGACCTGCTTTATGTGTATTTGCGCTTCCTC 300  
QY 319 ATCATCACCTACTCACTAGATGACAGTGGCCCTTCGGGGAGCTGCTCTCAAGCTGGTG 378  
DB 301 ATCATCACCTACTCACTAGATGACAGTGGCCCTTCGGGGAGCTGCTCTCAAGCTGGTG 360  
QY 379 CACTTCTCTGTCTATATATCAACCTTTACGGCAGCATCTCTGCTGTGACCTGCATCTCTGTG 438  
DB 361 CACTTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTGTGACCTGCATCTCTGTG 420  
QY 439 CACCACTTCTAGGTGTGTCACCCACTGTGTTGGTGGCCCTACCGGACCCCGAGGCAT 498  
DB 421 CACCACTTCTAGGTGTGTCACCCACTGTGTTGGTGGCCCTACCGGACCCCGAGGCAT 480  
QY 499 GCCTGGCTGGCAGCAGCACCCTGGCCCTGGTGGTCTCCAGCTGCTGCCACACTG 558  
DB 481 GCCTGGCTGGCAGCAGCACCCTGGCCCTGGTGGTCTCCAGCTGCTGCCACACTG 540  
QY 559 GCCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTGGTATGACATGACAGCCAA 618  
DB 541 GCCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGGTATGACATGACAGCCAA 600  
QY 619 GAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGTGACATTTGCTGCTTTCTTTC 678  
DB 601 GAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGTGACATTTGCTGCTTTCTTTC 660  
QY 679 CTCTTGGTCTATTTGGTGTGCTATTTACATGATGGTCTGAGGCTGATCAAGCAGAGGA 738  
DB 661 CTCTTGGTCTATTTGGTGTGCTATTTACATGATGGTCTGAGGCTGATCAAGCAGAGGA 720  
QY 739 GAACCTCATGAGGACAGGCAACACAGCCCGAGCGAGGTCCATCCGGACCATCTCTACTTGT 798  
DB 721 GAACCTCATGAGGACAGGCAACACAGCCCGAGCGAGGTCCATCCGGACCATCTCTACTTGT 780  
QY 799 GTGTGGCTCTTCACTCTGTGTTTGGTGGCTTCCCATATCACTGCTGCTTCTACTCTAC 858  
DB 781 GTGTGGCTCTTCACTCTGTGTTTGGTGGCTTCCCATATCACTGCTGCTTCTACTCTAC 840  
QY 859 CATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGGCAG-CCAGTGTGGCTTACA 917  
DB 841 CATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGGCAGCTTGTATGGCAGCTTGTGGCTTACA 900  
QY 918 AGATATGGAGGCTCTGTTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTACTTTCTTT 977  
DB 901 AGATATGGAGGCTCTGTTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTACTTTCTTT 960  
QY 978 CAAGGGGGCAAAATAGATGATGAGTCTCCAGAACTGAGGAGCAAGTGGGTGA 1037  
DB 961 CAAGGGGGCAAAATAGATGATGAGTCTCCAGAACTGAGGAGCAAGTGGGTGA 1020  
QY 1038 GCATCCAGCTGGGAGGAGAGATGCCAGGTTGAACAGATCT 1080  
DB 1021 GCATCCAGCTGGGAGGAGAGATGCCAGGTTGAACAGATCT 1063

RESULT 3  
AAD29667  
ID AAD29667 standard; cDNA; 1076 BP.  
XX  
AAD29667;  
XX AC  
XX AC  
DT 17-MAY-2002 (first entry)  
XX  
DE Human G-protein coupled receptor (GCREC-1) cDNA.  
XX  
KW Human; G-protein coupled receptor; GCREC-1; cell proliferative disorder;







angina pectoris; myocardial infarction; ulcer; inflammation; allergy;  
psychotic disorder; neurological disorder; anxiety; schizophrenia;  
manic depression; delirium; severe mental retardation; dyskinesia.

Homo sapiens.  
WO200185764-A2.  
15-NOV-2001.  
09-MAY-2001; 2001WO-EP05244.  
11-MAY-2000; 2000US-203582P.  
21-FEB-2001; 2001US-269857P.  
(FARB ) BAYER AG.

Ramakrishnan S;  
WPI; 2002-075242/10.  
New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders

Claim 1; Fig 4; 114pp; English.  
The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR. In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, osteoporosis, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies, psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic depression, delirium, severe mental retardation or dyskinesia). The present sequence is exon 1 of the P2Y-like GPCR of the invention.

Sequence 850 BP; 165 A; 262 C; 201 G; 222 T; 0 other;  
Query Match 78.68; Score 850; DB 24; Length 850;  
Best Local Similarity 100.08; Pred. No. 1.8e-221;  
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

85 ATGGAGAAGGTGGACATCAATACATCACAGGAACAGGTCTCTGCGAGTTCACAGAGAAG 144  
1 ATGGAGAAGGTGGACATCAATACATCACAGGAACAGGTCTCTGCGAGTTCACAGAGAAG 60  
145 TACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATCTAGGGCTGCCACTA 204  
61 TACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATCTAGGGCTGCCACTA 120  
205 ATGGCACTGCTTGTGGCACTCTGGGGCCCAACCAAGCGCTGGAGCTGGCCACCACC 264  
121 AATGGCACTGCTTGTGGCACTCTGGGGCCCAACCAAGCGCTGGAGCTGGCCACCACC 180  
265 TATCTGGTGAACCTCATGCTGGCCGACCTGCTTTATGCTATTGCTCCCTTCCTCATCATC 324  
181 TATCTGGTGAACCTCATGCTGGCCGACCTGCTTTATGCTATTGCTCCCTTCCTCATCATC 240

325 ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 384  
241 ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 300  
385 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTGACCTGCATCTCTGTGTGACCA 444  
301 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTGACCTGCATCTCTGTGTGACCA 360  
445 TTCTAGGTGTGTGCCACCCACTGTGTTGCTGCTTACCGGACCCGAGGAGCATGCTGCTGG 504  
361 TTCTAGGTGTGTGCCACCCACTGTGTTGCTGCTTACCGGACCCGAGGAGCATGCTGCTGG 420  
505 CTGGCACCAGCACACCTGGGCGCTGTGTGCTCTCCAGCTGCTGCCACACTGSCCTTC 564  
421 CTGGCACCAGCACACCTGGGCGCTGTGTGCTCTCCAGCTGCTGCCACACTGSCCTTC 480  
565 TCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGACATGACATGACCAAGCAAT 624  
481 TCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGACATGACCAAGCAAT 540  
625 TTTGATCGGCTTTTTCCTAGGCGATAGTTCTGACATTTGCTGGCTTTTCTTCCCTCCTT 684  
541 TTTGATCGGCTTTTTCCTAGGCGATAGTTCTGACATTTGCTGGCTTTTCTTCCCTCCTT 600  
685 GGTCAATTTTGTGTGCTATTCACATGATGGTCAGGAGCTGATCAAGCCAGAGGAGAACT 744  
601 GGTCAATTTTGTGTGCTATTCACATGATGGTCAGGAGCTGATCAAGCCAGAGGAGAACT 660  
745 CATGAGGACAGCAACACAGCCGAGGAGGATCCATCCGACCATCCTACTGCTGTGTGG 804  
661 CATGAGGACAGCAACACAGCCGAGGAGGATCCATCCGACCATCCTACTGCTGTGTGG 720  
805 CCTCTTACCCCTCTGTTTGTGCGCTTCCATATCACTCGCTCTTCTACCTCACCACATCTG 864  
721 CCTCTTACCCCTCTGTTTGTGCGCTTCCATATCACTCGCTCTTCTACCTCACCACATCTG 780  
865 CTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGTATGAGCAGCCAGTGTGGCTTCAAGATATG 924  
781 CTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGTATGAGCAGCCAGTGTGGCTTCAAGATATG 840  
925 GAGGCTCTCTG 934  
841 GAGGCTCTCTG 850  
RESULT 6  
AAH50975  
ID AAH50975 standard; DNA; 510 BP.  
XX AC AAH50975;  
XX DT 28-AUG-2001 (first entry)  
XX DE Human nPCR12 coding sequence.  
XX KW G protein-coupled receptor; nPCR; seven transmembrane receptor;  
XX KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
XX KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
XX KW cardiovascular disease; proliferative disorder; hormonal disorder;  
XX KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
XX KW attention deficit-hyperactivity disorder/attention deficit disorder;  
XX KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
XX KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
XX OS Homo sapiens.  
XX FN WO200136473-A2.  
XX PD 25-MAY-2001.  
XX PF 16-NOV-2000; 2000WO-US31581.



XX PR 16-NOV-1999; 99US-0165838.  
PR 17-NOV-1999; 99US-0166071.  
PR 19-NOV-1999; 99US-0166678.  
PR 28-DEC-1999; 99US-0173396.  
PR 22-FEB-2000; 2000US-0184129.  
PR 28-FEB-2000; 2000US-0185421.  
PR 28-FEB-2000; 2000US-0185554.  
PR 02-MAR-2000; 2000US-0186530.  
PR 03-MAR-2000; 2000US-0186811.  
PR 09-MAR-2000; 2000US-0188114.  
PR 17-MAR-2000; 2000US-0190310.  
PR 21-MAR-2000; 2000US-0190800.  
PR 20-APR-2000; 2000US-0198568.  
PR 02-MAY-2000; 2000US-0201190.  
PR 08-MAY-2000; 2000US-0203111.  
PR 25-MAY-2000; 2000US-0207094.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
PA Vogeli G, Wood LS, Hiebsch RR, Lind P, Slightom J;  
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
XX WPI; 2001-389826/41.  
DR P-PSDB; AAG80935.  
XX New G protein-coupled receptor (nGPCR-x) and its encoding  
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
XX Claim 4; Page 78; 261pp; English.  
CC The present invention relates to novel G protein-coupled receptors  
CC (nGPCR-x; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
CC sequence is the coding sequence for one such G protein-coupled receptor.  
CC nPCRs are also known as seven transmembrane receptors and function in  
CC signal transduction. The nGPCR coding sequences are useful for  
CC screening a human to diagnose a disorder affecting the brain or a genetic  
CC predisposition, specifically schizophrenia. nGPCR-x are useful for  
CC identifying compounds useful for treating schizophrenia. Detection of  
CC nGPCR in a sample is useful as a diagnostic tool for diseases or  
CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,  
CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular  
CC diseases, proliferative disorders and hormonal disorders. Modulators of  
CC nGPCR activity have the utility for treating neurological disorders,  
CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity  
CC disorder/attention deficit disorder), and neuronal disorders such as  
CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
CC Additional disorders include inflammatory conditions (e.g. Crohn's  
CC disease), rheumatoid arthritis, autoimmune disorders, cancers,  
CC respiratory ailments such as asthma, and inflammatory diseases e.g.  
CC inflammatory bowel disease.  
XX Sequence 510 BP; 92 A; 156 C; 126 G; 136 T; 0 other;  
SQ Query Match 47.1%; Score 510; DB 22; Length 510;  
Best Local Similarity 100.0%; Pred. No. 6.6e-129;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 247 TGGAGCTGTGCCACCACTATCTGTTGGTGAACCTGATGGTGGCCGACCTGTTATGTGCTA 306  
DB 1 TGGAGCTGTGCCACCACTATCTGTTGGTGAACCTGATGGTGGCCGACCTGTTATGTGCTA 60,  
QY 307 TTGGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 366  
DB 61 TTGGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120;  
QY 367 TGCAAGCTGTGCACTTCCCTGTTCTATATCAACCTTTACGGCAGCATCCTGCTGCTGACC 426  
DB 121 TGCAAGCTGTGCACTTCCCTGTTCTATATCAACCTTTACGGCAGCATCCTGCTGCTGACC 180  
QY 427 TGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTTGCTGCTGCCCTACCGG 486  
DB P-PSDB; AAU19205.

DB 181 TGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTTGCTGCCCTACCGG 240  
QY 487 ACCCGCAGGCATGCTGGCTGGGCACACACACCTGGGCCCTGGTGGTCTCCAGCTG 546  
DB 241 ACCCGCAGGCATGCTGGCTGGGCACACACACCTGGGCCCTGGTGGTCTCCAGCTG 300  
QY 547 CTGCCACACTGCGCTTCTCCACACGGACTACATCAATGCGCCAGATGATCTGATGAC 606  
DB 301 CTGCCACACTGCGCTTCTCCACACGGACTACATCAATGCGCCAGATGATCTGATGAC 360  
QY 607 ATGACAGCCCAAGAGAAATTTGATCGGCTTTTCCCTACGGCATAGTTCTGACATGCTCT 666  
DB 361 ATGACAGCCCAAGAGAAATTTGATCGGCTTTTCCCTACGGCATAGTTCTGACATGCTCT 420  
QY 667 GGCTTTCTTCCCTCTGCTGCTATTTGGTGTGCTATTCTACTGATGTCAGGAGCCTGAT 726  
DB 421 GGCTTTCTTCCCTCTGCTGCTATTTGGTGTGCTATTCTACTGATGTCAGGAGCCTGAT 480  
QY 727 CAAGCCAGAGGAGAACCTCATGAGGACAGG 756  
DB 481 CAAGCCAGAGGAGAACCTCATGAGGACAGG 510  
XX AAS30774 standard; cDNA; 681 BP.  
XX AAS30774;  
XX 04-DEC-2001 (first entry)  
XX Human cDNA encoding G protein-coupled receptor nGPCR-75.  
DE Human: G protein-coupled receptor; nGPCR-x; as; antiviral; analgesic;  
KW cytosolic; cardiac; antidiabetic; anorectic; hypotensive; hypertensive;  
KW antiparkinsonian; nootropic; neuroprotective; antidepressant;  
KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;  
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;  
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;  
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;  
KW schizophrenia; migraine; major depression; anxiety; mental disorder;  
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.  
OS Homo sapiens.  
XX WO200166750-A2.  
XX 13-SEP-2001.  
XX 08-MAR-2001; 2001WO-US07322.  
XX 08-MAR-2000; 2000US-0187581.  
XX 08-MAR-2000; 2000US-0187582.  
XX 08-MAR-2000; 2000US-0187714.  
XX 08-MAR-2000; 2000US-0187715.  
XX 08-MAR-2000; 2000US-0187825.  
XX 08-MAR-2000; 2000US-0187828.  
XX 08-MAR-2000; 2000US-0187829.  
XX 08-MAR-2000; 2000US-0187830.  
XX 08-MAR-2000; 2000US-0187833.  
XX 08-MAR-2000; 2000US-0187874.  
XX 08-MAR-2000; 2000US-0187930.  
XX 08-MAR-2000; 2000US-0188049.  
XX 08-MAR-2000; 2000US-0189294.  
XX 08-MAR-2000; 2000US-0189299.  
XX 08-MAR-2000; 2000US-0187928.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
PA Vogeli G, Wood LS;  
XX WPI; 2001-536778/59.  
DR P-PSDB; AAU19205.

XX Isolated nucleic acid molecules encoding G protein-coupled receptors  
PT termed nPCR-x, useful in the treatment and diagnosis of viral  
PT infections, cancers and mental disorders (e.g. Parkinson's disease and  
PT schizophrenia)  
XX  
XX Claim 4; Page 198; 336pp; English.  
PS  
XX The invention relates to novel isolated nucleic acid molecules encoding  
CC G protein-coupled receptors termed nPCR-x. nPCR-x polynucleotides,  
CC polypeptides, and modulators may be used in the treatment of diseases and  
CC conditions such as infections, such as viral infections caused by HIV-1  
CC (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and  
CC cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity,  
CC anorexia, hypotension, hypertension, myocardial infarction,  
CC atherosclerosis), Parkinson's disease, and psychotic and  
CC neurological disorders, including schizophrenia, migraine, major  
CC depression, anxiety, mental disorder, manic depression, and  
CC dyskinesias, such as Huntington's disease or Tourette's Syndrome  
CC and many other diseases and syndromes listed in the specification.  
CC nPCR-x polynucleotides and polypeptides, as well as nPCR-x  
CC modulators, may also be used in diagnostic assays for such diseases or  
CC conditions. The present sequence encodes a G protein-coupled  
CC receptor of the invention.

/ Sequence 681 BP; 163 A; 171 C; 171 G; 176 T; 0 other;

Query Match 36.6%; Score 396.4; DB 22; Length 681;  
Best Local Similarity 99.7%; Pred. No. 6.9e-98;  
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 685 GGTCAATTTGGTGTCTATCTACTGATGTCAGGAGCCTGATCAAGCCAGAGGAAACCT 744  
DB 681 GGTCAATTTGGTGTCTATCTACTGATGTCAGGAGCCTGATCAAGCCAGAGGAAACCT 622  
QY 745 CATGAGGACAGACACAGCCGAGCCAGGTCCATCCGACCATCTACTGGTGTGG 804  
DB 621 CATGAGGACAGACACAGCCGAGCCAGGTCCATCCGACCATCTACTGGTGTGG 562  
QY 805 CCTCTTCAACCCCTCTGTTTGTGCCCCCTTCATATCACTCGCTCCTTCTACCTCACCATCTG 864  
DB 561 CCTCTTCAACCCCTCTGTTTGTGCCCCCTTCATATCACTCGCTCCTTCTACCTCACCATCTG 502  
QY 865 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCGCAGCTGTGGCTACAGATATG 924  
DB 501 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCGCAGCTGTGGCTACAGATATG 442  
QY 925 GAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCCAGTCCGTACTTCTTTCAAGGGG 984  
DB 441 GAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCCAGTCCGTACTTCTTTCAAGGGG 382  
QY 985 GGCAGAAATAGAGTCAAGCTCCTCCAGAACTGAGGCGAGAACAGTTGGTGAGCATCCA 1044  
DB 381 GGCAGAAATAGAGTCAAGCTCCTCCAGAACTGAGGCGAGAACAGTTGGTGAGCATCCA 322  
QY 1045 GCTGGGAGGAAAGATGCCAGGTTTGACACATCTGG 1082  
DB 321 GCTGGGAGGAAAGATGCCAGGTTTGACACATCTGG 284

RESULT 8  
AAS30783

ID AAS30783 standard; cDNA; 585 BP.

XX AAS30783;

XX 04-DEC-2001 (first entry)

XX Human cDNA encoding G protein-coupled receptor nPCR-84.

XX Human; G protein-coupled receptor; nPCR-x; ss; antiviral; analgesic;  
KW cytosolic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive;  
KW antiparkinsonian; nootropic; neuroprotective; antidepressant;

KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;  
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;  
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;  
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;  
KW schizophrenia; migraine; major depression; anxiety; mental disorder;  
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.  
XX Homo sapiens.  
XX WO200166750-A2.  
XX 13-SEP-2001.  
XX 08-MAR-2001; 2001WO-US07322.  
PR 08-MAR-2000; 2000US-0187581.  
PR 08-MAR-2000; 2000US-0187582.  
PR 08-MAR-2000; 2000US-0187714.  
PR 08-MAR-2000; 2000US-0187715.  
PR 08-MAR-2000; 2000US-0187825.  
PR 08-MAR-2000; 2000US-0187828.  
PR 08-MAR-2000; 2000US-0187829.  
PR 08-MAR-2000; 2000US-0187830.  
PR 08-MAR-2000; 2000US-0187833.  
PR 08-MAR-2000; 2000US-0187874.  
PR 08-MAR-2000; 2000US-0187930.  
PR 08-MAR-2000; 2000US-0188049.  
PR 08-MAR-2000; 2000US-0189294.  
PR 08-MAR-2000; 2000US-0187929.  
PR 08-MAR-2000; 2000US-0187928.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX Vogeli G, Wood LS;  
XX WPI; 2001-536778/59.  
XX P-PSDB; AAU19214.  
XX Isolated nucleic acid molecules encoding G protein-coupled receptors  
PT termed nPCR-x, useful in the treatment and diagnosis of viral  
PT infections, cancers and mental disorders (e.g. Parkinson's disease and  
PT schizophrenia) -  
XX Claim 4; Page 201; 336pp; English.  
XX The invention relates to novel isolated nucleic acid molecules encoding  
CC G protein-coupled receptors termed nPCR-x. nPCR-x polynucleotides,  
CC polypeptides, and modulators may be used in the treatment of diseases and  
CC conditions such as infections, such as viral infections caused by HIV-1  
CC (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and  
CC cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity,  
CC anorexia, hypotension, hypertension, myocardial infarction,  
CC atherosclerosis), Parkinson's disease, and psychotic and  
CC neurological disorders, including schizophrenia, migraine, major  
CC depression, anxiety, mental disorder, manic depression, and  
CC dyskinesias, such as Huntington's disease or Tourette's Syndrome  
CC and many other diseases and syndromes listed in the specification.  
CC nPCR-x polynucleotides and polypeptides, as well as nPCR-x  
CC modulators, may also be used in diagnostic assays for such diseases or  
CC conditions. The present sequence encodes a G protein-coupled  
CC receptor of the invention.

Sequence 585 BP; 145 A; 142 C; 155 G; 143 T; 0 other;

Query Match 34.2%; Score 370; DB 22; Length 585;

Best Local Similarity 99.3%; Pred. No. 1e-90;

Matches 403; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 680 TCCTTGGTCATTTGGTGTGCTATTCTACTGATGGTCAGGAGCCTGATCAAGCCAGGAG 739

DB 1 TCCTTGGTCATTTGGTGTGCTATTCTACTGATGGTCAGGAGCCTGATCAAGCCAGGAG 60

QY 740 -AACCTCATGAGG-ACAGGCAACACAGCCGAGGCTCCATCC-GGACCATCTCTATG 796



```

Db 1038 CGCACCTCTACTACTCTCCCTCCGCT---CGCTGGACCTCAGCTGCCACACCTCAACGCC 1094
QY 902 GCAGGTGGCTGCTACAGATATGGAGGCTCTGGTGAGTGAGGAGCTGCCTCAACCCCA 961
Db 1095 ATCAACATGGCTTACAGGTATCCCGGCGCTGGCCAGTGCTAACAGTTGCTTGACCCC 1154
QY 962 GTCCTGACTTCTTTTCAAGG 982
Db 1155 GTGCTCTACTTCTCGCTGGG 1175

RESULT 10
AAD04981
ID AAD04981 standard; cDNA; 1163 BP.
AC AAD04981;
XX
DT 17-JUL-2001 (first entry)
XX
DE Human purinergic receptor P2Y2 cDNA.
DE
KW Human; purinergic receptor; P2Y2; antidepressant; vulnerary; hypotensive;
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;
KW irritable bowel disorder; reproductive system disorder; hypertension;
KW peripheral vascular disease; immune system disorder; chronic bronchitis;
KW premature ejaculation; asthma; neuromuscular disease; ss.
XX
OS Homo sapiens.
XX
FH
FT primer_bind Location/Qualifiers
FT complement (1..32)
FT /tag= a
FT /note= "Sense primer"
FT CDS
FT 15..1148
FT /tag= b
FT /product= "protein encoded by human P2Y2 cDNA"
FT /transl_except= (pos:786..788, aa:Phe)
FT misc_feature 787
FT /tag= c
FT /note= "Represented in the specification as F"
FT primer_bind 1131..1163
FT /tag= d
FT /bound_moiety= "Antisense primer"
XX
PN US6214581-B1.
XX
XX
XX 10-APR-2001.
XX
XX 13-NOV-1998; 98US-0191136.
XX
XX 16-JAN-1998; 98US-0071298.
XX 16-JAN-1998; 98US-0071669.
XX 16-JAN-1998; 98US-0008185.
XX 16-JAN-1998; 98US-0008526.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Lynch KJ, Burgard EC, Van Biesen T;
XX
XX WPI: 2001-315459/33.
XX
XX P-PSDB; AAE01114.
XX
XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor
XX polypeptide useful for identifying potentially therapeutic compounds
XX that modulate or otherwise interact with P2X containing receptors -
XX
XX Example 14; Flg 12A; 53pp; English.
XX
XX The present sequence is a cDNA encoding human P2Y2 receptor protein. P2X
XX receptors are ligand-gated ion channels while P2Y receptors operate
XX generally through a G-protein coupled system. P2X purinoreceptor drugs

```

CC are potential therapeutic agents in several disorders including central nervous system or peripheral nervous system conditions, e.g., epilepsy, pain, depression, neurodegenerative disorders, disorders of the skeletal muscle such as neuromuscular diseases, disorders of the reproductive system, asthma, peripheral vascular disease, hypertension, immune system disorders, irritable bowel disorder, premature ejaculation, cystic fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity of extracellular nucleotide triphosphates to regulate chloride secretion in human airway epithelia.

XX  
SQ Sequence 1163 BP; 177 A; 403 C; 332 G; 250 T; 1 other;

Query Match 20.7%; Score 223.6; DB 22; Length 1163;  
Best Local Similarity 56.1%; Pred. No. 1.le-50;  
Matches 483; Conservative 0; Mismatches 370; Indels 8; Gaps 3;

```

QY 127 TGCAGTTCCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCGCTACAGTATCATCTTT 186
Db 87 TGCCCGCTTCAAGGAGACTTCAAGTACGTGCTGCTGTCTCTACGGCTGGTGTGC 146
QY 187 ATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGC 246
Db 147 GTGCTTGGGCTGTGTCTGAACCGCTGGCGCTCTACATCTTCTTGGCGCTCAAGACC 206
QY 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGTGGTGCGGCGACCTGCTTTATGTGCTA 306
Db 207 TGGAAATGGCTGCCACCACTATATGTTCACCTGGCTGTCTGTGATGCATCTTATGCGGC 266
QY 307 T---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAGGTG 363
Db 267 TCCCTGGCGCTGCTGTCTATTACTACGCCCGCGGCGACCACTGGCCCTTTACGACGGTG 326
QY 364 CTCTCAAGCTGGTGCACCTTCTGTCTATATCAACCTTTACGGCAGCATCTCTGTCTG 423
Db 327 CTCTCAAGCTGGTGGCTTCTCTTCTACACCAACCTTTACTGCAGCATCTCTTCTCTC 386
QY 424 ACCTGCATCTCTGTGCACAGTTCTCTAGTGTGCGCACCCACTGTGTGCGTGCCTTAC 483
Db 387 ACCTGCATCAGCTGCACCGGTCTCTGGCGCTTCTACGACCTCTCGCTCCCTCGCTGG 446
QY 484 CGGACCCGCGAGCATGCTGTGGCGCACACGACACCTGTGGCGCTGTGTCTCTCCAC 543
Db 447 GCGCGGCGCGCTACGCTCGCGGGTGGCGGGCGCTGTGGGTGTGTGCTGGCTTGC 506
QY 544 CTGCTGCCCCACACTGGCTTCTCCACAGGACTACATCAATATGGCCAGATGATCTGGTAT 603
Db 507 CAGGCCCGCTGCTCTACTTGTCTACACACAGCGCGCGGGCGCGCTTAACCTGCCAC 566
QY 604 GACATGACCAAGCAAGAAATTTTCATCGGCTTTTGGCTACGGCATAGTTCTGACATTG 663
Db 567 GACACCTCGGCACCGAGCTCTTCAGCGCTTCTGGCGCTACAGCTCATGTCTGGGC 626
QY 664 TCTGGCTT--TCTTTCCCTCTCTGTGTCTATTTGGTGTGCTATTCATGATGTCTAGGAGC 721
Db 627 CTGCTCTTCGCGGTGCCCTTTGCCCTCATCTTGTCTGTAGTCTCATGCTCGCGCA 686
QY 722 CTGATCAAGCCAGAGGAAACCTCATAGGACAGCAACACAGCCCGAGCCAGGTCTTCATC 781
Db 687 CTGCTAAAGCCAGCCCTACGGGACCTTCGGCGCGCTTGCCTAGGGCCAAAGCAGTCCG 746
QY 782 CGGACCATCTACTGCTGTGGCTCTTACCCCTCTGTTTGTGCTCCCTTCCATATCACT 841
Db 747 CGCACCATCGCGGTGTGCTGTGCTGTCTTCTAGGACTGCCAGCTCTGTATGGCA 806
QY 842 CGCTCTCTTACTCACCATCTGCTTCTTCTCTAGGACTGCCAGCTCTGTATGGCA 901
Db 807 CGCACCTCTACTACTCTCTCG---TTCGCTGGACCTCAGCTGCCACACCTCAAGCC 863
QY 902 GCCAGTGTGGCTACAAGATATGGAGGCTCTGTGTGAGTGTGAGCAGTGTGCTCAACCCA 961
Db 864 ATCAACATGGCTTACAGGTATCCCGGCGCTGGCGCAGTGTCTAACAGTTGCTTGACCCC 923
QY 962 GTCTGTACTTCTTTCAGG 982

```

Db 924 GTGCTCTACTTCTCGCTGGG 944

RESULT 11  
AAS42837/c  
ID AAS42837 standard; cDNA; 551 BP.

XX AC AAS42837;

XX DT 18-DEC-2001 (first entry)

XX DE Human G Protein-Coupled Receptor: (GPCR) cDNA #32.

XX KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia; attention deficit disorder; anxiety; depression; bipolar disorder; ss; neurological disorder; Huntington's disease; dementia; obesity; anorexia; metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis; type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer; cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV; viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer; antidepressant; anorectic; PCR primer; gene therapy.

XX OS Homo sapiens.

XX PN WO200162797-A2.

XX PD 30-AUG-2001.

XX PF 23-FEB-2001; 2001WO-US05676.

XX PR 23-FEB-2000; 2000US-0184247.

XX PR 23-FEB-2000; 2000US-0184303.

XX PR 23-FEB-2000; 2000US-0184304.

XX PR 23-FEB-2000; 2000US-0184305.

XX PR 23-FEB-2000; 2000US-0184397.

XX PR 02-MAR-2000; 2000US-0186457.

XX PR 03-MAR-2000; 2000US-0186810.

XX PR 09-MAR-2000; 2000US-0188064.

XX PR 13-MAR-2000; 2000US-0188880.

XX PR 03-APR-2000; 2000US-0194344.

XX PR 23-JUN-2000; 2000US-0213861.

XX PR 11-JUL-2000; 2000US-0217369.

XX PR 11-JUL-2000; 2000US-0217370.

XX PR 14-JUL-2000; 2000US-0218337.

XX PR 20-JUL-2000; 2000US-0218492.

XX PA (PHRA ) PHARMACIA & UPJOHN CO.

XX PI Vogeli G, Wood LS, Parodi LA, Lind P;

XX DR WPI; 2001-570628/64.

XX DR P-PSDB; AAU25585.

XX PT New isolated nucleic acid encoding a new G-protein coupled receptor polypeptide for detecting receptor modulators that can treat mental disorders, such as schizophrenia, anxiety, depression, or obesity -

XX PS Claim 4; Page 82; 279pp; English.

XX CC Sequences AAS42806-AAS42926 represent cDNA molecules and PCR primers for cDNA molecules encoding human G-protein coupled receptor (GPCR) polypeptides. The protein and DNA sequences of the invention can be used to identify compounds which bind to GPCR polypeptides and in screening for compounds that modulate GPCR activity. By screening a human subject for the presence of mutations in GPCR DNA, a GPCR-related disorder or a genetic predisposition can be diagnosed. The sequences can also be used for treatment and prevention of mental disorders such as schizophrenia, attention deficit disorder, anxiety, depression, dementia and bipolar disorder, neurological disorders such as Huntington's disease, Parkinson's disease and Tourette's syndrome, metabolic disorders such as obesity, anorexia and type 2 diabetes, cardiovascular disorders such as thrombosis, myocardial infarction, cardiomyopathy and atherosclerosis,

CC viral infections caused by HIV and cancers.  
XX  
SQ Sequence 551 BP; 143 A; 146 C; 130 G; 132 T; 0 other;

Query Match 20.5%; Score 222; DB 22; Length 551;  
Best Local Similarity 100.0%; Pred. No. 2.1e-50;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTGCTTCCAGGGAAGCAGAGCGGGAGCGCTCGTGGAGCTCTG 60  
|||||  
Db 222 ATGCTGTCCATTGCTTCCAGGGAAGCAGAGCGGGAGCGCTCGTGGAGCTCTG 163  
|||||

QY 61 CTCCTGGAGGAGCCCTCCCGGACATGGAGAAGTGGACATGAATACATCAGGAACAA 120  
|||||

Db 162 CTCCTGGAGGAGCCCTCCCGGACATGGAGAAGTGGACATGAATACATCAGGAACAA 103  
|||||

QY 121 GGTCTGTCCAGTTCACAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACATATC 180  
|||||

Db 102 GGTCTGTCCAGTTCACAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACATATC 43  
|||||

QY 181 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTGTGG 222  
|||||

Db 42 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTGTGG 1  
|||||

RESULT 12  
AAQ88134  
ID AAQ88134 standard; cDNA; 1842 BP.

XX AC AAQ88134;

XX DT 29-NOV-1995 (first entry)

XX DE Human P20 receptor gene.

XX KW Epithelial mucosa; mucus; cystic fibrosis; asthma; ss;

XX KW chronic bronchitis.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 57..1181

XX FT /\*tag= a

XX PN WO9510538-A.

XX PD 20-APR-1995.

XX PF 04-OCT-1994; 94WO-US11260.

XX PR 15-OCT-1993; 93US-0138137.

XX PA (UMOR ) UNIV MISSOURI.

XX PA (UYNC-) UNIV NORTH CAROLINA.

XX PI Boucher RC, Erb LJ, Harden TK, Lustig KD, Parr CE;

XX PI Sullivan DM, Turner JT, Weisman GA;

XX DR WPI; 1995-169967/22.

XX DR P-PSDB; AAR72457.

XX DNA encoding human P20 receptor and null cells expressing the receptors - for stimulating or inhibiting growth of cultures of mammalian cells, and for treating diseases of airway epithelial mucosa, e.g. asthma.

XX Claim 3; Page 34; 47pp; English.

XX The sequence is that of the human P20 receptor gene. The gene may be used to treat diseases of epithelial mucosal surfaces by enhancing or inhibiting mucus prodn. The gene may be used to treat cystic fibrosis, asthma and chronic bronchitis.  
CC See also AAQ88135-40.

XX	SQ	Sequence 1842 BP; 364 A; 557 C; 528 G; 393 T; 0 other;
XX	Query Match	18.7%; Score 202.2; DB 16; Length 1842;
XX	Best Local Similarity	56.2%; Pred. No. 9e-45;
XX	Matches 484; Conservative 0; Mismatches 363; Indels 14; Gaps 5;	
QY	127	TGCCAGTCTCTCAGAGAAGTACAAAGTCTACCTCTCCCTGGCTACAGTATCATCTTT 186
DB	129	TGCCCTTCAAGGAGGACTTGAAGTACGTGCTGCTGCTCTACGGCTGTGTGC 188
QY	187	ATCCTAGGCTGCCACTAAATGGCACTCTCTTTGGGCACTCTCTGGGCAACACCAAGCCG 246
DB	189	GTGCTTGGGCTGTGTCTGAACCGCCCTGGGCTCTACATCTCTTGTGGCGCTCAAGACC 248
QY	247	TGGAGCTGTGCCACCACTATCTGTGTGAACCTGTGTGGCGACCTCTTTATCTGCTA 306
DB	249	TGGAATGGCTCACCAACATATATGTTCACCTGGGTGTGTGATGACACTGTATCGGGC 308
QY	307	T---TGCCCTTCTCATCATCACCTACTACTAGATGACAGTGGCCCTTCGGGGAGCTG 363
DB	309	TCCCTGCGCTGTCTATTACTATACGCGCGCGGCACTGGCCCTTCAGCACGGTG 368
QY	364	CTCTCAAGCTGGTCACTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTG 423
DB	369	CTCTCAAGCTGGTGGCTCTCTCTTCTACACCACTTTACTGCAGCATCTCTTCTCTC 428
QY	424	ACCTGCATCTCTGTGCACCACTCTCTAGTGTGTGCCACCACTGTGTGCTGCCCTTAC 483
DB	429	ACCTGCATCAGCTGCACCGGTGTCTGGCGCTCTTACGACCTCTCGCTCCCTGGCTGG 488
QY	484	CGGACCGGAGCATGCTGCTGGTGGGCAACAGCACCTCTGGGCGCTGGTGTCTCCAGC 543
DB	489	GGCGGGGCGGCTAGCTGCGCGGGTGGCGGGCGGTGGG---TGTGTGTGCTGGCC 545
QY	544	CTGCTGCCCACTGCGCTCTCCACACGAGTACATCAATGCGGCACTGATCTGCTAT 603
DB	546	TCCAGGCGCGCTCTACTTTGTACACGAGCGGGCGGGCGGCTTAACCTGCGCAC 605
QY	604	GACATGACCAAGCAAGAAATTTGATCGGCTTTTGGCTAGGCGATAGTTCTGACATTG 663
DB	606	GACACCTCGGCACCGAGCTCTTACGCGCTTCGTGGCTACAGCTCAGTCACTGCTGGC 665
QY	664	TCTGCTT--TCTTTCCCTCTTGTGCTATTTTGGTGTGCTATCTACTGATGCTAGGAGC 721
DB	666	CTGCTCTTTCGGGCTGCCCTTTGCGCTCATCTTGTCTTACGTCTCATGCTCGGCGA 725
QY	722	CTGATCAAGCCAGAGGAACTCTCATGAGGACGACCAACAGCCGAGGAGTCCATC 781
DB	726	CTGCTAAGCCAGGCTACGGGACCTCGGGCGGCTGCTAGGGCCCAAGGCAAGTCCCTG 785
QY	782	CGGACCACTCTACTGCTGTGTGGCTCTTACGCTCTCTTTTGTGCCCTTCCATATCACT 841
DB	786	CGCACCATCGCGGTGTGTGCTGTCTTCCGCTCTCTCTCTCTCTCCATCCAGCTCAAC 845
QY	842	CGCTCCTCTACTACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCA 901
DB	846	CGCACCTCTACTACTCTCTCGCT--CGCTGGACCTCAGCTGCGCACACCTCAACGCC 902
QY	902	GCCAGTGTGGCTCAACATATGAGGCTCTGTGTGAGTGTGAGGCTGCTCAACCCA 961
DB	903	ATCAACATGGGCTCAAGGT---TACCGGCTGGCCAGTGTCAACAGTTGCTTGCACCC 959
QY	962	GTCTGTACTTCTTTCAAGG 982
DB	960	GTGCTTACTTCTGCTGGG 980
XX	RESULT 13	
XX	AAD32937	
XX	ID	AAD32937 standard; DNA; 2138 BP.
XX	AC	AAD32937;

XX	01-JUL-2002	(first entry)
XX	DT	Mus musculus GPCR ATP-receptor P2U1 gene.
XX	DE	
XX	DE	Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor; genetic therapy; pharmacological; genetic disease; neuropsychological; neurological; psychotic illness; gene; nephrotropic; gynaecological; psychostimulant; ATP-receptor P2U1; ds.
XX	OS	Mus musculus.
XX	Key	Location/Qualifiers
XX	CDS	270..1391
FT		/*tag= a
FT		/product= "ATP-receptor P2U1 protein"
FT	misc_feature	1569..1570
FT		/*tag= b
FT		/note= "Illegible in the specification"
XX	WO200203793-A2.	
XX	17-JAN-2002.	
XX	10-JUL-2001;	2001WO-US21923.
XX	10-JUL-2000;	2000US-217058P.
XX	10-JUL-2000;	2000US-217179P.
XX	10-JUL-2000;	2000US-217223P.
XX	10-JUL-2000;	2000US-217253P.
XX	10-JUL-2000;	2000US-217255P.
XX	10-JUL-2000;	2000US-217256P.
XX	10-JUL-2000;	2000US-217257P.
XX	11-JUL-2000;	2000US-217347P.
XX	11-JUL-2000;	2000US-217629P.
XX	12-JUL-2000;	2000US-217537P.
XX	12-JUL-2000;	2000US-218069P.
XX	12-JUL-2000;	2000US-218074P.
XX	12-JUL-2000;	2000US-218358P.
XX	27-JUL-2000;	2000US-221483P.
XX	07-AUG-2000;	2000US-223120P.
XX	07-AUG-2000;	2000US-223122P.
XX	26-OCT-2000;	2000US-243958P.
XX	15-NOV-2000;	2000US-249408P.
XX	20-NOV-2000;	2000US-252299P.
XX	16-JAN-2001;	2001US-262113P.
XX	16-JAN-2001;	2001US-262205P.
XX	(DELT-)	DELTA GEN INC.
XX	Allen KD,	Brennan TJ;
PI		
XX	WPI;	2002-164574/21.
XX	P-PSDB;	RAE20604.
XX	Novel non-human transgenic animal, especially transgenic mice useful for identifying an agent that modulates expression or function of target gene, comprises disruptions in target G protein-coupled receptor gene	
XX	Disclosure; Fig 19; 101pp; English.	
XX	The invention relates to a non-human transgenic animal having targeted G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene 5-HT-2B gene, chemokine receptor 9A gene, glucocorticoid-induced receptor gene, orphan GPR10 (UHR)-1 gene, orphan GPR14 gene, orphan GPR15 gene, beta chemokine receptor (E01) gene, endothelial differentiation GPCR3 (EDG3) gene, ATP receptor P2U1 gene or adenosine 3 receptor gene. The transgenic animal is useful for identifying an agent that modulates the expression or function of the target gene, for identifying an agent that ameliorates a phenotype associated with a disruption in the target gene. A transgenic construct is useful for producing a transgenic animal, preferably a transgenic mouse. The transgenic animal is useful for	

CC testing the efficacy of proposed genetic and pharmacological therapies  
CC for human genetic diseases, such as neurological, neuropsychological or  
CC psychotic illnesses. The transgenic animal is also useful as models for  
CC diseases, disorders or conditions associated with phenotypes relating to  
CC a disruption in a target, and to identify pharmaceuticals, therapies,  
CC drugs and interventions which may be effective in treating a disease or  
CC other phenotypic characteristics of the animal. An agent which modulates  
CC the expression of the target gene is useful as a therapeutic for treating  
CC conditions associated with a disruption of the target gene. The present  
CC sequence is mouse GPCR ATP-receptor P2U1 gene.  
XX

SQ Sequence 2138 BP; 423 A; 623 C; 606 G; 480 T; 6 other;

Query Match 18.7%; Score 202.2; DB 24; Length 2138;  
Best Local Similarity 54.6%; Pred. No. 9.6e-45;  
Matches 470; Conservative 0; Mismatches 383; Indels 8; Gaps 3;  
QY 127 TGGCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186  
DB 342 TGTCTGCTTCAACGAGGACTTCAAGTACGTCTGTTGCCGCTGTCTATGGCGTGGTGTGC 401  
QY 187 ATCTAGGCTGCCACTAATAGCACTGTCTGTGGCACTCTCTGGGCGCCAAACCAAGCGC 246  
DB 402 GTCTCGGGTGTGCTGACGCTGTGGCTCTATATCTTCTATGCGCGCTCAAAACC 461  
QY 247 TGGAGCTGTGCCACCACTATCTGTGGTGAACCTGATGGTGGCCGCACTGCTTTATGTGCTA 306  
DB 462 TGAACGCCCTCCACCACTACATGTTTACCTGGCAGTTTCGCACTCTCTACGAGCGC 521  
QY 307 T----TCCCTTCCTCATCACTACCTACTACTAGATGACAGGTGGCCCTTCGGGAGCTG 363  
DB 522 TCCCTCGCGCTGTGGTATTATTAACGCGCGGGTGACCACTGGCCATTTAGCAGGTTG 581  
QY 364 CTCTGAAGCTGTGCATCTCTCTTCTATCAACCTTTACGGCAGCATCTCTGCTGCTG 423  
DB 582 CTCTGAAGCTGTGGTGTCTCTCTTCTACCAACCTCTACTGCGACATCTCTCTCTCCTC 641  
QY 424 ACCTGATCTCTGTGACCACTAGTGTGTGGTGGCCACCACTGTGTGCTGCGCTTAC 483  
DB 642 ACCTGATCAGCTGACCGTGGCTGGGAGTCTTGGCCCTCTGCACTCCCTGCGTGG 701  
QY 484 CGAACCGGAGGATGCTGTGGTGGGACAGCAGCACTGGCCCTGGTGGTCTCTCCAG 543  
DB 702 GGCGGGCGCGTTATGCCCGCGGGTGGTGGCTGTGTGGTGTGTGGTGTGGCTGTC 761  
QY 544 CTGCTGCCACACTGCTCTTCTCCCACTGAGTACATCAATGCGCAGATGATCTGGTAT 603  
DB 762 CAGGACCGCTGCTCTACTTCTACCACTGCGGGGAGCCGATCATCTTGCAT 821  
QY 604 GACATGACCAAGCAAGAGATTTTGTATCGGCTTTTTCCTACGCGCATAGTCTG--ACAT 661  
DB 822 GACACCTCGCGCGAGAGCTCTTTAGCCATTTTGTGGCTTACAGCTCCGTCATGCTGGT 881  
QY 662 TGTCTGGCTTTCTTCCCTCTGTGCTATTTGGTGTGCTATTCATGATGGTCAAGAGC 721  
DB 882 CTGCTTTTCTGTGCGCTTTTCCGTAATCTGCTGTGTATGCTGTGTATGGCGACGGG 941  
QY 722 CTGATCAAGCAGGAGACCTCATGAGCAGCAACACACCCGCGAGCCAGCTCCATC 781  
DB 942 CTGCTCAACCGGCTTATGGGACCAAGAGGTCTGCTCGGGCGCAAGCAAGTCTGTG 1001  
QY 782 CGGACCATCTACTGTGTGGCGCTTTTCCACCTCTGTTTGTGGCTTTCCCATATCACT 841  
DB 1002 CGACCATTTGCTGTGCTGCTTCTGCGCTCTGCTTTCTGCTTTCCACGTCAGC 1061  
QY 842 CGCTCTCTTACCTACCACTGCTGCTTTCTGCTTTCTGAGCAGTGGCAGCTCTTTGATGGCA 901  
DB 1062 CGCACCTCTTACTACTCTCTTCCGAT---CACTTGAACCTCAGCTGCCACACCTCAAGCC 1118  
QY 902 GCCAGTGTGGCTTACAGATATGGAGCCCTCTGTGAGTGTGAGCAGCTCCCTCAACCCA 961  
DB 1119 ATCAACATGGCATATAGATCACCGCGCGCTGGCGAGCGCCCAACAGTGTCTCTTGACCG 1178

QY 962 GTCCTGCTACTTTCTTTTCAAGG 982  
DB 1179 GACTCTACTTCTGCGAGG 1199  
RESULT 14  
AAD32944  
ID AAD32944 standard; DNA; 2138 BP.  
XX AAD32944;  
AC AAD32944;  
XX  
DT 01-JUL-2002 (first entry)  
XX Mus musculus GPCR ATP-receptor P2U1 gene targeting construct.  
DE DE  
KW Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor;  
KW genetic therapy; pharmacological; genetic disease; neuropsychological;  
KW neurological; psychotic illness; nephrotropic; gynaecological;  
KW psychostimulant; ATP-receptor P2U1; ds.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 266..350  
FT /tag= a  
FT /note= "Sequence flanking Neo insert in targeting  
FT construct"  
FT misc\_feature 351..636  
FT /tag= b  
FT /note= "Sequence deleted in targeting construct"  
FT misc\_feature 637..941  
FT /tag= c  
FT /note= "Sequence flanking Neo insert in targeting  
FT construct"  
XX  
PN WO200203793-A2.  
XX  
PD 17-JAN-2002.  
XX  
PF 10-JUL-2001; 2001WO-US21923.  
XX  
PR 10-JUL-2000; 2000US-217058P.  
PR 10-JUL-2000; 2000US-217179P.  
PR 10-JUL-2000; 2000US-217223P.  
PR 10-JUL-2000; 2000US-217253P.  
PR 10-JUL-2000; 2000US-217255P.  
PR 10-JUL-2000; 2000US-217256P.  
PR 10-JUL-2000; 2000US-217257P.  
PR 11-JUL-2000; 2000US-217347P.  
PR 11-JUL-2000; 2000US-217629P.  
PR 12-JUL-2000; 2000US-217537P.  
PR 12-JUL-2000; 2000US-218069P.  
PR 12-JUL-2000; 2000US-218074P.  
PR 12-JUL-2000; 2000US-218358P.  
PR 27-JUL-2000; 2000US-221483P.  
PR 07-AUG-2000; 2000US-223120P.  
PR 07-AUG-2000; 2000US-223122P.  
PR 26-OCT-2000; 2000US-243958P.  
PR 15-NOV-2000; 2000US-249408P.  
PR 20-NOV-2000; 2000US-252299P.  
PR 16-JAN-2001; 2001US-262113P.  
PR 16-JAN-2001; 2001US-262205P.  
XX (DELT-) DELTAGEN INC.  
XX  
PI Allen KD, Brennan TJ;  
XX WPI; 2002-164574/21.  
XX  
PT Novel non-human transgenic animal, especially transgenic mice useful  
PT for identifying an agent that modulates expression or function of  
PT target gene, comprises disruptions in target G protein-coupled receptor  
PT gene



XX PS Disclosure: Fig 20A; 101pp; English.

CC The invention relates to a non-human transgenic animal having targeted

CC G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene

CC 5-HT<sub>2A</sub> gene, chemokine receptor 9A gene, glucocorticoid-induced receptor

CC gene, orphan GPR10 (UHR)-1 gene, orphan GPR14 gene, orphan GPR15 gene,

CC beta chemokine receptor (E01) gene, endothelial differentiation GPCR3

CC (EDG3) gene, ATP receptor P2U1 gene or adenosine 3 receptor gene. The

CC transgenic animal is useful for identifying an agent that modulates the

CC expression or function of the target gene, for identifying an agent that

CC ameliorates a phenotype associated with a disruption in the target gene.

CC A transgenic construct is useful for producing a transgenic animal,

CC preferably a transgenic mouse. The transgenic animal is useful for

CC testing the efficacy of proposed genetic and pharmacological therapies

CC for human genetic diseases, such as neurological, neuropsychological or

CC psychotic illnesses. The transgenic animal is also useful as models for

CC diseases, disorders or conditions associated with phenotypes relating to

CC a disruption in a target, and to identify pharmaceuticals, therapies,

CC drugs and interventions which may be effective in treating a disease or

CC other phenotypic characteristics of the animal. An agent which modulates

CC the expression of the target gene is useful as a therapeutic for treating

CC conditions associated with a disruption of the target gene. The present

CC sequence is mouse GPCR ATP-receptor P2U1 gene targeting construct.

Sequence 2138 BP; 423 A; 623 C; 608 G; 480 T; 4 other;

Query Match 18.7%; Score 202.2; DB 24; Length 2138;  
Best Local Similarity 54.6%; Pred. No. 9.6e-45;  
Matches 470; Conservative 0; Mismatches 383; Indels 8; Gaps 3;

Qy 127 TGCCAGTTCTCAGAGAACTACAAGCAAGCTACCTCTCCCTGGCCACAGTATCATCTTT 186

Db 342 TGTGCTTCAACGAGGACTTCAAGTACGTGCTGTGGCCGCTGCTATGGCTGGTGTGC 401

Qy 187 ATCTAGGCTGCCACATAATGGCACTCTCTTGGGCACTCTCTGGGGCCAAACGAGCC 246

Db 402 GTGCTCGGGTTGTGCTGAAGCTGTGGCTCTCTATATCTTCTATGCGCCCTCAAAACC 461

Qy 247 TGGAGCTGTGCACACCATCTATCTGTGAACATGTGTGGCGGACCTCTTTATGTGCTA 306

Db 462 TGGAAAGCTCCACCACTATCTTCACTGTTCACCTGGCAGTTTCGGACTCTCTACGACGC 521

Qy 307 T---TGCCCTTCTCATCATCACCCTACCTACCTAGATGACAGTGGCCCTCGGGAGCTG 363

Db 522 TCCCTGCGCTGTGGTTTATTACTACGCCGGGTGACCACTGGCCATTTAGCACGGTG 581

Qy 364 CTCGTGCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTGCTG 423

Db 582 CTCGTCAAGCTGGTGCGTTTCTCTTCTACCAACCTCTACTGCAGCATCTCTTCTCCTC 641

Qy 424 ACCTGCACTCTGTGCAACAGTTCTCTAGGTGTGTGCCACCACTGTGTGCTGCGCCCTAC 483

Db 642 ACCTGCACTACAGCTGCAACGGTGTGCTGGAGTCTGCGCCCTCTGCACTCCTCGGTGG 701

Qy 484 CGGACCCGCAAGCATGCTGCTGGCCACCAAGCACTGCGCCCTGGTGGTCTCTCCAG 543

Db 702 GGCGGGCCGCTTATGCGCCGCGGTGCTGCGGTGTGTGGTGTGCTGCTGCGCTGCG 761

Qy 544 CTGCTGCCACACTGGCTTCTCCACAGGCACTACATCAATGGCCAGATCATCTGCTAT 603

Db 762 CAGGACCCGCTGCTACTCTGCTCACCACAGCGTGGGGGAACCCCGGATCACTTGGCCAT 821

Qy 604 GACATGACACCAAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTTCTG--ACAT 661

Db 822 GACACCTGGCCGAGAGCTCTTAGCCATTTTGTGGCTTACAGCTCGCTATGCTGGGT 881

Qy 662 TGCTGGCTTTCTTCCCTCTCTGGTCAATTTTGTGTGCTATTTCACTGATGGTCAGGAGC 721

Db 882 CTGCTTTTGTGTGCGCTTTTCCGTAATCTGCTGTGCTTACGTGCTTATGCGCCAGGG 941

Qy 722 CTGATCAGCCAGGAGAACTCATGAGGACAGCAACAGCCCGGACAGGTCCATC 781

Db 942 CTGCTCAAAACCGGCTTATGGGACCACAGAGGTCTGCTCGGGCCAAAGCGAAGTCTGTG 1001

Qy 782 CGGACCATCTCTACTGCTGTGGCTCTTACCCCTCTCTTTTGTGCTTCCATATCACT 841

Db 1002 CGACCATTTGCTGGTACTGGCGCTCTTGGCCCTCTCTTCTGCTTCCACGCTCAG 1061

Qy 842 CGCTCTTCTTACCTACCATCTGCTTCTTCTCTCAGGACTGCCAGCTCTTTGATGCA 901

Db 1062 CGACCCCTCTACTACTCTCTCCGAT---CACTTGACCTCAGCTGCCACACCTCAACGCC 1118

Qy 902 GCCAGTGTGGCTACAAATATGGAGGCTCTGGTGTAGTGTGAGCAGCTGCTCAACCCA 961

Db 1119 ATCAACATGGCATATAAGATCACCGCGCTGGCCAGCGCAACAGTGTGTTGACCCG 1178

Qy 962 GTCCTGTACTTCTTTTCAAGG 982

Db 1179 GTACTCTACTTCTCTGGCAGG 1199

RESULT 15

AAT74321

ID AAT74321 standard; DNA; 1429 BP.

XX AC AAT74321;

XX DT 31-MAR-1998 (first entry)

XX DE Human P2Y4 receptor genomic DNA.

XX KW Receptor; P2Y4; pyrimidine binding; uridine triphosphate; UTP; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 181..1278

XX FT /\*tag= a

XX FT /\*product= P2Y4 receptor

XX PN W09719170-A1.

XX PD 29-MAY-1997.

XX PF 21-NOV-1996; 96WO-BE00123.

XX PR 21-NOV-1995; 95EP-0870124.

XX PA (EURO-) EUROSREEN SA.

XX PI Boeynaems J, Communi D, Parmentier M, Pirotton S;

XX DR WPI; 1997-402177/37.

XX DR P-PSDB; AAW23606.

XX PT Receptor having preference for pyrimidine over purine nucleotide(s).

XX PT - especially uridine triphosphate, agonist and antagonists of which

XX PS are useful in treatment of cystic fibrosis

XX PS Claim 14; Figure 1; 56pp; English.

XX CC This sequence encodes a novel human P2 receptor, P2Y4, which has a

CC preference for pyrimidine binding, especially uridine triphosphate.

CC This receptor could be used to screen for novel drugs which

CC specifically bind to it. Transgenic animals could be used to determine

CC the physiological effects of expressing varying levels of the receptor

CC or to identify novel agonists or antagonists. The agonists and

CC antagonists of human P2Y4 may be used, e.g., in treatment of cystic

CC fibrosis.

SQ Sequence 1429 BP; 242 A; 451 C; 379 G; 357 T; 0 other;

Query Match 16.5%; Score 178.6; DB 18; Length 1429;  
Best Local Similarity 54.0%; Pred. No. 2.2e-38;  
Matches 481; Conservative 0; Mismatches 394; Indels 16; Gaps 5;









```
Query Match      18.8%; Score 203.8; DB 1; Length 1842;
Best Local Similarity 56.3%; Pred. No. 1.7e-46;
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;

Qy 127 TGCCAGTTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 TGCCGCTTCAACGAGGACTTCAAGTACGTGTGCTGCTGCTGTCTACGGGTGCTGTC 188

Qy 187 ATCCTAGGCTGCCACTAAAGTCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 GTGCTGGGCTGTGTGTAACGCGTCTACATCTTCTGTGGCGCTCAAGACC 248

Qy 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGTGTGGGCGCTCTACATCTTCTATCTGCTA 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 TGGATGCGTCCACCACATATATGTTCCACCTGGCTGTCTGATGCACTGTATCGCGCC 308

Qy 307 T---TGCCCTTCTCATCATCCTACTACTAGATGATGAGTGGCCCTTCGGGGAGCTG 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 TCCCTGGCGCTGCTGGTCTATTACTACCGCGCGGCGGCGGCTTTCAGCAGCGTG 368

Qy 364 CTCGCAAGCTGGTGCATCTTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTG 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 CTCGCAAGCTGGTGGCTTCTCTTCTACCAACCTTTACTGACGATCTCTTCTCTC 428

Qy 424 ACCTGCATCTCTGTGCACCACTTCTTAGTGTGTGCCACCACTGTGTGCTGCTGCTAC 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 2  
US-08-444-581B-1  
; Sequence 1, Application US/08444581B  
; Patent No. 5607836  
; GENERAL INFORMATION:  
; APPLICANT: Boucher, Richard C.  
; APPLICANT: Weisman, Gary A.

```
; APPLICANT: Turner, John T.  
; APPLICANT: Harden, Thomas K.  
; APPLICANT: Parr, Claude E.  
; APPLICANT: Sullivan, Daniel M.  
; APPLICANT: Erb, Laura  
; APPLICANT: Lustig, Kevin D.  
; TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and  
; TITLE OF INVENTION: Null Cells Expressing P2U Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bell, Seltzer, Park & Gibson  
; STREET: Post Office Drawer 34009  
; CITY: Charlotte  
; STATE: No. 5607836th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,581B  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/442,134  
; FILING DATE: 16-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sibley, Kenneth D.  
; REGISTRATION NUMBER: 31,665  
; REFERENCE/DOCKET NUMBER: 5470-71A  
; TELEPHONE: 919-420-2200  
; TELEFAX: 919-881-3175  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1842 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 57..1181  
; US-08-444-581B-1
```

```
Query Match      18.8%; Score 203.8; DB 1; Length 1842;  
Best Local Similarity 56.3%; Pred. No. 1.7e-46;  
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;  
Qy 127 TGCCAGTTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186  
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 129 TGCCGCTTCAACGAGGACTTCAAGTACGTGTGCTGCTGCTGTCTACGGGTGCTGTC 188  
Qy 187 ATCCTAGGCTGCCACTAAAGTCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 246  
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 189 GTGCTGGGCTGTGTGTAACGCGTCTACATCTTCTGTGGCGCTCAAGACC 248  
Qy 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGTGTGGGCGCTCTACATCTTCTATCTGCTA 306  
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 249 TGGATGCGTCCACCACATATATGTTCCACCTGGCTGTCTGATGCACTGTATCGCGCC 308  
Qy 307 T---TGCCCTTCTCATCATCCTACTACTAGATGATGAGTGGCCCTTCGGGGAGCTG 363  
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 309 TCCCTGGCGCTGCTGGTCTATTACTACCGCGCGGCGGCGGCTTTCAGCAGCGTG 368  
Qy 364 CTCGCAAGCTGGTGCATCTTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTG 423  
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 369 CTCGCAAGCTGGTGGCTTCTCTTCTACCAACCTTTACTGACGATCTCTTCTCTC 428  
Qy 424 ACCTGCATCTCTGTGCACCACTTCTTAGTGTGTGCCACCACTGTGTGCTGCTGCTAC 483  
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 429 ACCTGATCAGCGTGCACCGGTGCTGGGGCTTTACGACCTCTGCGCTCCCTGCGCTGG 488
QY 484 CGGACCCGAGGATGCTGGTGGGACACGACCACTGGGCGCTGGTGGTCTCCAG 543
Db 489 GCGCGGCGGCTACGCTCGCGGTGGCGGGCGGTGGG---TGTGGTGGTGGCC 545
QY 544 CTGCTGCCACACTGCGCTTCTCCACAGGACTATCAATGAGGCGAGATGCTGGTAT 603
Db 546 TGCCAGGCGCGGTGCTCTACTTTGTACACAGCGCGCGGGCGCTAACCTGCCAC 605
QY 604 GACATGACGAGCAAGAAATTTGATCGCTTTTGGCTTACGCGATAGTTCTGACATG 663
Db 606 GACACCTCGGACCGGCTCTACGCGCTTCTGGGCGCTGCTGAGCTCAGCTCATGCTGGG 665
QY 664 TCTGGCTT---TCTTTCCCTTGGTCAATTTGGTGTCTATTCTACTGATGCTGAGG 721
Db 666 CTGCTCTTCCGCGTGGCTTTGCGGCTTCTGCTGTCTGTCTGCTCATGCTGCTGCGGA 725
QY 722 CTGATCAAGCGAGGAGAACTCTCATGAGGACAGGCAACAGCGGAGGCGAGCTGCATC 781
Db 726 CTGCTAAAGCGGAGGCTACGGGACCTCGGGGCGCTCCCTAGGGCGGAGGAGTCCG 785
QY 782 CGGACCATCTACTGCTGTGGTGGCTCTTACCGCTCTTACCGCTCTTCCATATCACT 841
Db 786 CGACCATCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 845
QY 842 CGCTCCTTACTCCTACCATCTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTG 901
Db 846 CGACCCCTACTACTCTCTTCTTCAAGG 982
QY 902 GCGAGTGGCTTACAGATAGGAGGCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 961
Db 903 ATCAACATGCGCTTACAGGT---TACCGGCTGGCGAGTGTAAACAGTTGGCTTGAC 959
QY 962 GTCCCTACTTCTTCTTCAAGG 982
Db 960 GTGCTACTTCTGCTGGG 980

RESULT 3
US-08-446-088A-1
; Sequence 1, Application US/08446088A
; Patent No. 5691156
GENERAL INFORMATION:
APPLICANT: Boucher, Richard C.
APPLICANT: Weisman, Gary A.
APPLICANT: Turner, John T.
APPLICANT: Harden, Thomas K.
APPLICANT: Parr, Claude E.
APPLICANT: Sullivan, Daniel M.
APPLICANT: Erb, Laura
APPLICANT: Lustig, Kevin D.
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
TITLE OF INVENTION: Null Cells Expressing P2U Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: No. 5691156th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,088A
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D. Sibley
```

```
;
;
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-71C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1842 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1181
; US-08-446-088A-1

Query Match 18.8%; Score 203.8; DB 1; Length 1842;
Best Local Similarity 56.3%; Pred. No. 1.7e-46;
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;

QY 127 TGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186
Db 129 TGCCGCTTCAAGAGGACTTCAAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188
QY 187 ATCCTAGGCTGCCAATAAATGGCACTGCTTGTGGCACTCTCTGGGCGGCAAAACAGCGC 246
Db 189 GTGCTTGGGCTGTGCTGAACGCGGTGGGCTCTACATCTTCTTGTGGCGCTCAAGACC 248
QY 247 TGGAGTGTGGCAACCACTATCTGTGGAACCTGATGGTGGCGGACCTGCTTTATGCTA 306
Db 249 TGGAAATGCGCTCCACCACATATATGTTCCACCTGGCTGTGTCTGATGACCTGTATGCG 308
QY 307 T---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTG 363
Db 309 TCCTGCGCGTGTGCTGCTTACTAGCGCGGCGGACACTGGGCGCTTCACAGCGGTG 368
QY 364 CTCTGCAAGCTGTGCACTTCTCTTATATCAAGCTTTTACGGCAGCATCTCTGCTGCTG 423
Db 369 CTCTGCAAGCTGTGCGCTTCTCTTCTACACCAACTTTTACTGACGATCTCTTCTCCTC 428
QY 424 ACCTGATCTCTGTGCACAGTCTCTAGTGTGTGTGCACACCACTGTGTGCTGCGCTAC 483
Db 429 ACCTGATCAGCGTGCACGCGTCTTGGGCGTCTTACGACCTCTGCGCTCCCTGCGCTG 488
QY 484 CGGACCCGAGGATGCTGGTGGCGCACGACCACTGGGCGCTGGTGGTGGTGGTGGTGG 543
Db 489 GCGCGGCGCGCTTACCTGCGGCGGCGGCGGCGGTGGG---TGTGGTGGTGGGCG 545
QY 544 CTGCTGCCACACTGCGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGCTAT 603
Db 546 TGCCAGGCGCGCTGCTCTTCTTGTACACAGCGCGCGCGCGCTTAACCTGCCAC 605
QY 604 GACATGACGAGCAAGAAATTTGATGCGCTTTTGGCTTACGCGCATAGTCTTGTGACAT 663
Db 606 GACACCTCGGCAACCGGAGCTCTTACGCGCTTCTGGGCGCTTACAGCTCAGTCTGCGG 665
QY 664 TCTGGCTT---TCTTTCCCTTGGTCAATTTGGTGTGCTTATTCATCTGATGCTGAGG 721
Db 666 CTGCTCTTCCGCGTGGCTTTGCGGCTTCTGCTGTCTGTCTGTCTGTCTGTCTGCTGCG 725
QY 722 CTGATCAAGCGAGGAGAACTCTCATGAGGACAGGCAACAGCGGAGGCGAGCTGCATC 781
Db 726 CTGCTAAAGCGGAGGCTACGGGACCTCGGGGCGCTCCCTAGGGCGGAGGAGTCCG 785
QY 782 CGGACCATCTACTGCTGTGGTGGCTCTTACCGCTCTTACCGCTCTTCCATATCACT 841
Db 786 CGACCATCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 845
QY 842 CGCTCCTTCTACTCCTACCATCTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTG 901
Db 846 CGACCCCTACTACTCTCTTCTTCAAGG 982
```



Db 831 GGCTTTTGAGGGGCTACAAAGCAGCGGGCGTTTGCCAGTGCACAGCGGTGCTGGA 890

QY 958 CCAGTCCTGTACT 971

Db 891 CCCATCCTCTTCT 904

RESULT 5

US-08-513-974B-379  
; Sequence 379, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ogi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236356  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189274  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189273  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189272  
; FILING DATE: 11-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Resnick, David S.  
; REGISTRATION NUMBER: 34,235

; REFERENCE/DOCKET NUMBER: 45753

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO.: 379;

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1023 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 37..1020

US-08-513-974B-379

Query Match

Best Local Similarity 52.0%; Score 164.8; DB 3; Length 1023;

Matches 444; Conservative 0; Mismatches 402; Indels 8; Gaps 3;

QY 126 CTGCCAGTTCTCAGAGAAGTACAAAGTCTACCTCTCCCTGGCCTACAGTATCATCTT 185

Db 87 CTGTGTCTACCGGAGAACTTCAAGCAACTGCTGCTGCCACCTGTGTATTCGGGTGCT 146

QY 186 TATCCTTAGGCTGCCACTAAATGGCCTGTCTTTGTGGCCTCTCCCTGGGCGCAAGCG 245

Db 147 GCGGGCTGGCTGCGCTGAACATCTGTGTCATTACCCAGATCTGCACGTCCCGCGGC 206

QY 246 CTGGAGCTGTGCCACCACTATCTGTGAACCTGTGGTGGCGGACCTGCTTTATGTGCT 305

Db 207 CTTGACCCGCGCGGCTGTACACCTTAAACCTTGTCTGGCTGACCTGCTATATGCTG 266

QY 306 AT---TGCCCTTCTCATCATCACTACTTCTACTAGTACAGTGGCCCTTCGGGGAGCT 362

Db 267 CTCCTGCCCCCTGCTCATCTACAACCTATGCCAAGGTGATCACTGGCCCTTTTGGCACT 326

QY 363 GCTCTCAAGCTGTGCACTTCTCTGTATATCAACCTTTACGGCAGCATCTGCTGCT 422

Db 327 CGCTGCGGCTGTGCTCGCTTCTCTTATGCCAACCTGCACGGCAGCATCTCTCTTCT 386

QY 423 GACCTGCATCTGTGCACCACTTCTTAGTGTGTGCCACCACTGTGTGCTGGCTGCCTA 482

Db 387 CACCTGCATCAGCTTCCAGCGCTACCTGGGCATCTGCCACCCGCTGGCCCTTGGCA 446

QY 483 CC---GGACCCGAGCATGCTGGCTGGGCACCAAGCAGCAGCAGCAGCAGCAGCAGC 539

Db 447 ACCTGGGGCGCGGGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506

QY 540 CCAGCTGCTGCCACACTGGCCCTTCTCCACAGGACTACATCAATGGCAGATGATCTG 599

Db 507 CCAGTGCCTGCCACAGCCATCTTGGCTGCCACAGGATCCAGGTAACCGCACTGTCTG 566

QY 600 GTATGACATGACCAAGCAAGAAATTTTGATCGCTTTTGGCTACGCACTGCTGCTGAC 659

Db 567 CTATGACCTCAGCCCGCTGCTGGCCACCCACTATATGCCCTATGCTGCTGCTGCTG 626

QY 660 ATTGCTCTGGCTTCT---TTCCTCTCTGGTCAATTTGGTGTCTTATCACTGATGTGCT 717

Db 627 TGTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686

QY 718 GAGCCTGATCAAGCCAGAGAGAACTCTATGAGACAGGCAACACAGCCCGAGCAGGTC 777

Db 687 CCGCTGTGCGGCGAGGATGGCCCGGAGAGCCTGTGGCCAGAGCGGCTGGCAAGGC 746

QY 778 CATCCGGACCATCTACTGCTGTGGCTTCTTCCACCTCTTCCACCTCTGTTTGGCTTCCATAT 837

Db 747 GCGCCGATGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806

QY 838 CACTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 897

Db 807 CACCAAGACAGCCTACCTGCGAGTGGCTCGACGCGGGGCTGCCCTGCACTGTATTGGA 866

QY 898 GGCAGCCAGTGTGGCCTTACAAGATATGGAGGCGCTCTGCTGCTGCTGCTGCTGCTGCT 957

```
Db 867 GGCCTTTGACGGCCTACAAAGGACCGCGCGTTTCCAGTGCCAAACAGCGTGTCTGGA 926
Qy 958 CCCAGTCTCTGTACT 971
Db 927 CCCCATCTCTTCT 940

RESULT 6
US-08-459-046-1
; Sequence 1, Application US/08459046
; Patent No. 6008039
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 Hillview Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF-0038 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Placenta
; CLONE: 179696
; US-08-459-046-1

; Query Match 14.5% Score 156.8; DB 3; Length 984;
; Best Local Similarity 51.4%; Pred. No. 1e-33;
; Matches 439; Conservative 0; Mismatches 407; Indels 8; Gaps 3;

Qy 126 CTCGACGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGATCATCTT 185
Db 51 CTGTGTTCTACCGCAGAGACTTCAAGCAACTGCTGCCACCTGTGTATTCGGCGGTGCT 110
Qy 186 TATCTAGGCGTCCGACCAATATGGCACTGTCTTGTGGCACTCTCTGGGCGCAACCAAGCG 245
Db 111 GCGCGCTGCCCTCCGCTGGAACATCTGTGTCAATTACCCAGATCTGCAGTCCCGCGCGGC 170
Qy 246 CTGAGAGTGTGCCACCACTATCTGGTGAACCTGATGTCGCGCACCTGCTTTATGTGCT 305
Db 171 CTGACCGCGACGCGGTGTACACCTTAAACCTTGTCTGCTGCTGACCTGCTATATGCGTG 230
Qy 306 AT---TGCCCTTCTCATCATCATCACTACTAGATGACAGGTGGCCCTTCGGGGAGCT 362
Db 231 CTCCTGCCCCCTGCTCATCTACAACTATGCGCAAGGTGATCACTGGCCCTTTGGCGACTT 290
```

```
Qy 363 GCTCTGCAAGTGGTGGCACTTCCTGTTCTATATCAACCTTTACGGCAGCATCCTCTCTGCT 422
Db 291 GCGCTGCGCGCTGGTCCGCTTCCTCTTCTATGCCAACTGCACGGGAGGATCCTCTTCTCT 350
Qy 423 GACCTGCATCTCTGTGCACCAAGTTCCCTAGCTGTGGCCACCACTGTCTTGGCTGCCCTA 482
Db 351 CACCTGCATCAGCTTCCAGCGCTACCTGGGCATCTGCCACCGCTGGCCCTTGGCCCAAA 410
Qy 483 CC---GGACCCGCGAGGATGCTGGTGGCAACCAAGCAACCTTCGGTGGTGGTCTCTCT 539
Db 411 ACGTGGGCGCGCGGCTGCTGGCTAGTGTGTAGCGCTGTGGCTGGCGTGACACAA 470
Qy 540 CCAGTGTGTCGCCACACTGGGCTTCTCCACAGGACTACATCAATAGGCCAGATGATCTG 599
Db 471 CCAGTGCCTGGCCACAGCATCTTCGCTGCCACAGGCATCCAGCGTAAACCGCACTGTCTG 530
Qy 600 GTATGACATGACAGGCCAAGAGAAATTTTGTATCGGCTTTTGTGCTACGCGCATAGTTCTGAC 659
Db 531 TTATGACCTCAGCCGCGCTGCGCTGGCCACCACTATATGCCCTATGGGATGGCTCTCAC 590
Qy 660 ATTGTCTGGCTTTCT--TTCCCTCTCCTTGGTCAATTTGGTGTGCTATTTCACCTGATGGTCAG 717
Db 591 TGTATCGGCTTCTGCTGCGCTTTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
Qy 718 GAGCTGTATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACAGCCGCGAGCGGCTG 777
Db 651 CCGCTGTGCGCGCAGGATGGCGGAGAGCTGTGCGCCAGGAGCTGTGCGCCAGGAGCGCTG 710
Qy 778 CATCGGACCATCTACTGTGTGTGGCTCTTCCACCTCTGTTTGTGCGCTTCCATAT 837
Db 711 GCGCGCATGGCGGTGGTGGCTGCTGCTTTGGCATCAGCTTCTGCTGCTTTTTCACAT 770
Qy 838 CACTCGCTCTTCTACCTCAACCATCTGCTTTCTCTCAGGACTGCCAGCTTCTGAT 897
Db 771 CACCAAGACACGCTACCTGGCAGTGGCTGCGACGCGCGGCTGCCCTGCATGTATTGGA 830
Qy 898 GCGACCCAGTGTGGCTACAAGATATGAGGCGCTCTGTGTAGTGTGAGCAGCTGCTCAA 957
Db 831 GCGCTTTGACGCGCTTACAAGGACGCGCGCGCTTTGCCAGTGCACACAGCGTGTGGA 890
Qy 958 CCCAGTCTCTGACT 971
Db 891 CCCCATCTCTTCT 904

RESULT 7
US-08-513-974B-41
; Sequence 41, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
```





PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01599  
FILING DATE: 10-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 370:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1020 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..1011  
US-08-513-974B-370

very Match 13.8%; Score 148.8; DB 3; Length 1020;  
st Local Similarity 50.8%; Pred. No. 1.7e-31;  
Matches 434; Conservative 0; Mismatches 412; Indels 8; Gaps 3;  
Qy 126 CTGCGAGTCTCAGAGTACAGGAGTCTACCTCTCCCTGCGCCACAGATATCATCTT 185  
Db 78 CTGGCTACCGTACGAGTATCAAGCGACTGCTGTAACCCCGGTATATCTCGGTGCT 137  
Qy 186 TATCTAGGCTGCCATTAATGCACTGTCTTGTGGCACTCTTGGGCGCAACCAAGCG 245  
Db 138 GGTGGTGGCGCTGCACATGAACATCTGCTCATTTGCCAGATCTGCGCATGCCCGCGAC 197  
Qy 246 CTGGAGTGTGCCACCACTATCTGTTGGAACCTGATGTGGCGACCTGCTTTATGTGCT 305  
Db 198 CCTGACCGTTCGCTGTGTACACCTGAACTGGCAGCTGCGGACCTGATGATGCTG 257  
Qy 306 AT---TGCCCTTCTCATCATCTACTACTAGATACAGGTTGGCCCTTTCGGGAGCT 362  
Db 258 TTCACTACCCCTACTTATCTATACTACGCCAGAGGGGACCACTGGCCCTTCGGAGACT 317  
Qy 363 GCTCTGCAAGCTGTGTCACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTCT 422

Db 318 GCGCTGCGCGCTTTGTAGCCTTCTCTTCTATGCAATCTACATGCGCAGCATCTGTTCT 377  
Qy 423 GACCTGCATCTCTGTGCACCAAGTTCCCTAGGTGTGTGCGCACCCACTGTGTGCTGCCCTA 482  
Db 378 CACCTGCATTTAGCTTCCAGCGCTACCTGGGCATCTGCCACCCCTGGCTTCTGTCGACAA 437  
Qy 483 CC---GGACCCGCGAGGATGCTTGGCTGGGCACCAAGCAACCACTGGGCGCTTGGTGTCT 539  
Db 438 GCGTGGAGTGGCGCTGCTGCTTGGTAGTGTGTGGAGTCTGTGGCTGTGCTGTGACAGC 497  
Qy 540 CCAGCTGTGCCACACACTGGCCCTTCTCCACACAGGACTACATCAATGCGCAGAGATCTG 599  
Db 498 CCAGTGGCTGCCACCGCAGTCTTTTGTGTCACAGGATCCAGCGCAACCCGACCTGTGTG 557  
Qy 600 GTATGACATGACCAAGCAAGAGAAATTTTGTGCGCTTTTTCCTACGCGCATAGTTCTGAC 659  
Db 558 CTAGCACTGAGCCCAACCCATCTCTACTCTGCTACCTGCGCTATGATGGCCCTCAC 617  
Qy 660 ATTCTGTGGCT--TTCTTTCCCTCCTTGGTCAATTTGGTGTGCTATTCACTGATGGTCA 717  
Db 618 GGTCACTGGCTTCTTGTGCTGCTTCTATAGCCTTACTGGCTTGTATTGTCGATGGCCG 677  
Qy 718 GAGCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCGAGTC 777  
Db 678 CCGCTGTGTGCGCAGGATGCGCCAGCAGGTCTGTGTGCGCCCAAGAGCGCGCAAGGC 737  
Qy 778 CATCCGACCATCTACTTGTGTGCTGCTTCAACCTCTGTTTGTGCGCTTCCATAT 837  
Db 738 GGCTGATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797  
Qy 838 CACTGCTGCTTCTACCTCAACCTCTGCTTCTGCTTCTTCTCAGGACTGCCAGCTTGTGAT 897  
Db 798 CACCAAGACAGCTTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857  
Qy 898 GCGAGCAGTGTGCGCTACAGATATGAGGCGCTCTGGTGTGCTGAGTGTGAGCAGCTGCC 957  
Db 858 GACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917  
Qy 958 CCCAGTCTCTTACT 971  
Db 918 CCCATCTCTTCT 931  
RESULT 9  
US-08-097-938-1  
Sequence 1, Application US/08097938  
Patent No. 5629174  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT CL40 RECEPTOR AND ITS  
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,938  
FILING DATE: 26-JUL-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 22803-20006.00



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,840

FILING DATE: 25-JAN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/390,301

FILING DATE: 25-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: ADLER, REID G.

REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 2803-0006.20

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1475 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 232..1416

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 232

US-08-472-840-1

Query Match 8.5%; Score 92; DB 1; Length 1475;  
Best Local Similarity 56.5%; Pred. No. 7.8e-16;  
Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

Qy 134 TCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 193  
Db 431 TCACGGGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 490  
Qy 194 GCCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGCAACCAAGCGCTGGAGCT 253  
Db 491 GTTGGCCCAAGTAAATGGCACTGTCTTGTGGCACTCTCTGGGCAACCAAGCGCTGGAGCT 550  
Qy 254 GTGCCACCACTATCTGTGTGAACCTGTGTGGCCAGCTCTCTGTGTGTGTGTGTGTGTGTGT 313  
Db 551 CGCGCGTATTTACATGGCAACCTGTGTGGCCAGCTCTCTGTGTGTGTGTGTGTGTGTGTGT 610  
Qy 314 TCCTCATCATCCTACTACTAGTAGTATG---ACAGTGGCCCTTCGGGAGCTGCTCTGCA 370  
Db 611 CCCTGAAGATCTCTTACCACTTACATGGCAACCTGTGTGGCCAGCTCTCTGTGTGTGTGTGT 670  
Qy 371 AGCTGGTGCACCTCTCTGT 430  
Db 671 AGGTGCTCATTTGGCTTTTCTATGTATACATGTATGTCTCTCTCTCTCTCTCTCTCTCTCT 730  
Qy 431 TCTCTGTGCACAGTCTCTAGT 470  
Db 731 TCAGGCGTCAGAGGTACTGGGT 770

RESULT 12  
US-08-476-976-1  
Sequence 1, Application US/08476976  
Patent No. 5874400  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20008-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,976  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1475 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 232..1416  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 232  
US-08-476-976-1

Query Match 8.5%; Score 92; DB 2; Length 1475;  
Best Local Similarity 56.5%; Pred. No. 7.8e-16;  
Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

Qy 134 TCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 193  
Db 431 TCACGGGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 490  
Qy 194 GCCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGCAACCAAGCGCTGGAGCT 253  
Db 491 GTTGGCCCAAGTAAATGGCACTGTCTTGTGGCACTCTCTGGGCAACCAAGCGCTGGAGCT 550  
Qy 254 GTGCCACCACTATCTGTGTGAACCTGTGTGGCCAGCTCTCTGTGTGTGTGTGTGTGTGTGT 313  
Db 551 CGCGCGTATTTACATGGCAACCTGTGTGGCCAGCTCTCTGTGTGTGTGTGTGTGTGTGTGT 610  
Qy 314 TCCTCATCATCCTACTACTAGTAGTATG---ACAGTGGCCCTTCGGGAGCTGCTCTGCA 370  
Db 611 CCCTGAAGATCTCTTACCACTTACATGGCAACCTGTGTGGCCAGCTCTCTGTGTGTGTGTGT 670  
Qy 371 AGCTGGTGCACCTCTCTGT 430  
Db 671 AGGTGCTCATTTGGCTTTTCTATGTATACATGTATGTCTCTCTCTCTCTCTCTCTCTCTCT 730  
Qy 431 TCTCTGTGCACAGTCTCTAGT 470  
Db 731 TCAGGCGTCAGAGGTACTGGGT 770

RESULT 13  
US-08-474-410-1  
Sequence 1, Application US/08474410  
Patent No. 6043212  
GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN  
 APPLICANT: SCARBOROUGH, ROBERT M.  
 TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
 AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
 NUMBER OF SEQUENCES: 63  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20006-1812  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/474,410  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/390,301  
 FILING DATE: 25-JAN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ADLER, REID G.  
 REGISTRATION NUMBER: 30,988  
 REFERENCE/DOCKET NUMBER: 2803-0006.20  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 TELEX: 90-4030  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1475 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 232..1416  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 232  
 -08-474-410-1

Query Match	8.5%	Score 92	DB 3	Length 1475
Best Local Similarity	56.5%	Pred. No. 7.8e-16		
Matches 192	Conservative 0	Mismatches 145	Indels 3	Gaps 1
Qy 134	TCCTCAGAGAGTACAACGAAGTCTACCTCTCCCTGGCGCTACAGTATCATCTTTTATCCTAG	193		
Db 431	TCACCGGGAAGCTGACCACCGGTCTTTCTCCGGTCGCTACATTATTGTGTTGTGATG	490		
Qy 194	GGGTGCCACCTAAATGTGCACCTCTTGTGGCACTCCTCTGGGCGAAACCAAGCGCTGGAGCT	253		
Db 491	GTTTGCCCAAGTAATGGCATGCCCTCTGGATCTTCTTCTCCGAACGAAGAACAACACC	550		
Qy 254	GTGCCACCACTATCTCGGTGAACATGATGGTGGCCGACCTGCTTTATGTGCTATTGGCCT	313		
Db 551	CCGCGGTGATTATACATGGCCAAACCTGGCCTTGGCCGACCTCCTCTCTCATCTGGTTCC	610		
Qy 314	TCCTCATCATCACCCTACTACTAGATG- --ACAGGTGGCCCTTCGGGAGCTGCTCTGCA	370		
Db 611	CCCTGAAGATCTCCTACCACTACATGGCAACAACCTGGGTCTACGGGAGGCCCTGTGCA	670		
Qy 371	AGTGTGTGCACCTTCTCTGTTCTATATCAACCTTTTACGGCAGCATCCTGCTGCTGACCTGCA	430		
Db 671	AGGTGCTCATTGGCTTTTCTATGATGATGATGCTCCATCCTCTTCATGACCTGCC	730		
Qy 431	TCTCTGTGCACCACTTCTCTAGGTGTGTGCCACCACCTGTG	470		
Db 731	TCAGCGTGCAGAGTACTGGGTGATCGTGAACCCATGGG	770		

RESULT 14

US-08-486-673B-1

US-08-486-673B-1, Application US/08486673B

Patent No. 6297026

GENERAL INFORMATION:

APPLICANT: Sundelin, Johan

APPLICANT: Scarborough, Robert M.

TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor

FILE REFERENCE: 44461-5006-08-US

CURRENT APPLICATION NUMBER: US/08/486,673B

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: US 08/097,938

PRIOR FILING DATE: 1993-07-26

PRIOR APPLICATION NUMBER: PCT/US94/08536

PRIOR FILING DATE: 1994-07-26

NUMBER OF SEQ ID NOS: 63

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 1475

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: CDS

LOCATION: (232)..(1416)

OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein

OTHER INFORMATION: sequences

US-08-486-673B-1

Query Match 8.5%; Score 92; DB 4; Length 1475;

Best Local Similarity 56.5%; Pred. No. 7.8e-16;

Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 134 TCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGGCTACAGTATCATCTTTATCCTAG 193

DB 431 TCACCGGAAGCTGACCACCGTCTTCTCCGGTCGCTACATATATGTGTTGTGATTG 490

QY 194 GCGTGGCACTAAATGGGCACCTGTCTTGTGGCAGCTCCCTGGGCGCAACCAACGCGTGGAGCT 253

DB 491 GTTTGGCCAGTAATGGCATGGCCCTCTGTGATCTTCTCTTTCCGAACGAAGAAGAACACC 550

QY 254 GTGCCACCACTATCTGGTGAACCTGATGTGCGCGGACCTGCTTTATGTCTATTGCCCT 313

DB 551 CCGCGGTGATTTACATGGCCACACTGGCCCTTGGCGGACCTCTCTCTGTCTATGGTTCC 610

QY 314 TCCTCATCATCACCTACTACTAGATG---ACAGGTGGCCCTTCGGGGAGCTGCTCTGCA 370

DB 611 CCCTGAAGATCTCCTACCACCTACATGCAACAACATGGGTCTACGGGGAGGCCCTGTGCA 670

QY 371 AGCTGGTGCACTTCTCTTTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGCA 430

DB 671 AGTGTCCTCATGGCTTTTCTATGTGTAACATGATTGTCTCCATCCTCTTTCATGACCTGCC 730

QY 431 TCTCTGTGCACCACTCTCTAGGTGTGTGCCACCCACCTGTG 470

DB 731 TCACCGTGCAGAGTACTGGGTGATCGTGAACCCCATGGG 770

RESULT 15

US-08-476-000-60

US-08-476-000-60, Application US/08476000

Patent No. 5716789

GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN

APPLICANT: SCARBOROUGH, ROBERT M.

TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

CITY: Washington

STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,000  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 73..1269  
US-08-476-000-60

Query Match 8.5% Score 92; DB 1; Length 2732;  
Best Local Similarity 56.5% Pred. No. 1e-15;  
Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;  
Qy 134 TCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCGGCTACAGTATCATCTTTATCCTAG 193  
Db 284 TCACCGGAAGCTGACCACGCTCTTCTCCGGTCGTACATTATTGTTGTTGATTG 343  
Qy 194 GGCTGCCACTAAATGGCACTCTCTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCT 253  
Db 344 GTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCTTTCCGAACGAAGAAACACC 403  
Qy 254 GTGCCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTATGCTATTGGCCT 313  
Db 404 CCGCGGTGATTACATGCCCAACCTGGCCCTGGCGGACCTCTCTGTCTATCTGGTTCC 463  
Qy 314 TCCTCATCATCACCCTACTACTAGATG---ACAGGTGGCCCTTCGGGGAGCTGCTCTGCA 370  
Db 464 CCCTGAAGATCTCTACCACTACATGCAACAACCTGGGCTACGGGGAGGCCCTGTGCA 523  
Qy 371 AGCTGGTGCACCTCTCTCTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGCA 430  
Db 524 AGGTGCTCATGGCTTTTCTATGTAACATGATGCTCCATCTCTCTTACCTGACCCTGCC 583  
Qy 431 TCTCTGTGCACCACTTCTAGTGTGTGCCACCACTGTG 470  
Db 584 TCACGTGCAGAGGTACTGGGTGATGCTGAACCCCATGGG 623

Search completed: May 26, 2003, 22:54:17  
Job time : 73 secs

GenCore version: 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003. CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 21:24:13 ; Search time 199 Seconds  
(without alignments)  
7179.599 Million cell updates/sec

Title: US-10-023-586b-1

Perfect score: 1082

Sequence: 1 atgctgtccatttgcctcc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1006	93.0	1017	9	US-09-885-453-3
2	510	47.1	510	9	US-09-782-974C-13
3	222	20.5	551	9	US-09-791-932-32
4	178.6	16.5	1429	9	US-09-077-173A-1
5	97.8	9.0	1543	9	US-09-891-138A-1
6	97.4	9.0	1773	9	US-10-112-599A-3
7	95.6	8.8	1892	9	US-09-900-699A-1
8	95.2	8.8	422	9	US-09-779-679-23
9	89	8.2	1526	10	US-09-728-952-40
10	88.4	8.2	1020	10	US-09-788-133-1
11	87.4	8.1	1679	9	US-09-104-063-5
12	86	7.9	1313	10	US-09-728-422-1
13	85.8	7.9	348	10	US-09-864-761-21546
14	83.6	7.7	1829	10	US-09-823-114-7
15	83.6	7.7	2218	10	US-09-214-904-3
16	83.6	7.7	2219	9	US-10-112-599A-1
17	83.2	7.7	1164	9	US-10-228-264-3
18	83.2	7.7	1365	7	US-08-899-112-29
19	82.6	7.6	2693	9	US-09-850-948-1
					Sequence 3, Appli
					Sequence 13, Appl
					Sequence 32, Appl
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 23, Appl
					Sequence 40, Appl
					Sequence 1, Appl
					Sequence 5, Appli
					Sequence 1, Appl
					Sequence 21546,
					Sequence 7, Appl
					Sequence 3, Appl
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 29, Appl
					Sequence 1, Appli

20	82	7.6	1805	10	US-09-823-114-18	Sequence 18, Appl
21	82	7.6	1829	9	US-09-905-186A-9	Sequence 9, Appli
22	82	7.6	1829	9	US-09-905-186A-10	Sequence 10, Appl
23	82	7.6	1829	9	US-09-905-186A-11	Sequence 11, Appl
24	82	7.6	2534	9	US-10-087-345A-22	Sequence 22, Appl
25	81.8	7.6	2602	9	US-09-905-186A-1	Sequence 1, Appli
26	80.8	7.5	1102	9	US-09-870-759-143	Sequence 143, App
27	80.8	7.5	1670	10	US-09-880-107-2143	Sequence 2143, Ap
28	80.8	7.5	1737	9	US-09-104-063-3	Sequence 3, Appli
29	80.4	7.4	1829	9	US-09-905-186A-7	Sequence 7, Appli
30	80.4	7.4	1829	9	US-09-905-186A-8	Sequence 8, Appli
31	79.6	7.4	2427	10	US-09-254-783A-2	Sequence 2, Appli
32	79.6	7.4	2427	12	US-10-152-058-2	Sequence 2, Appli
33	78.4	7.2	442	10	US-09-864-761-4808	Sequence 4808, Ap
34	77.6	7.2	1107	9	US-10-212-980-3	Sequence 3, Appli
35	77.6	7.2	1107	9	US-10-228-264-1	Sequence 1, Appli
36	77.6	7.2	1417	9	US-10-007-132-3	Sequence 3, Appli
37	77.6	7.2	1417	9	US-09-771-287-3	Sequence 3, Appli
38	77.4	7.2	1586	10	US-09-104-792-1	Sequence 1, Appli
39	77	7.1	1098	9	US-09-850-948-5	Sequence 5, Appli
40	77	7.1	1697	9	US-10-109-533A-1	Sequence 1, Appli
41	76.4	7.1	1404	10	US-09-805-628-3	Sequence 3, Appli
42	76.4	7.1	1419	10	US-09-805-628-1	Sequence 1, Appli
43	76.2	7.0	318	10	US-09-728-445-777	Sequence 777, App
44	76.2	7.0	1014	9	US-10-023-775B-1	Sequence 1, Appli
45	76.2	7.0	1014	9	US-10-270-144-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-885-453-3  
; Sequence 3, Application US/09885453  
; Publication No. US20030088080A1  
; GENERAL INFORMATION:  
; APPLICANT: Communi, Didier  
; TITLE OF INVENTION: RECEPTOR GPCRxl0  
; FILE REFERENCE: 9409/2082  
; CURRENT APPLICATION NUMBER: US/09/885,453  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/885,453  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1017  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DNA sequence  
; LOCATION: (1)..(1017)  
; OTHER INFORMATION: GPCRxl6 DNA sequence  
US-09-885-453-3

Query Match	93.0%;	Score	1006;	DB	9;	Length	1017;
Best Local Similarity	99.9%;	Pred. No.	1.4e-306;				
Matches	1017;	Conservative	0;	Mismatches	0;	Indels	1;
Gaps	1;						
QY	1	ATGCTGTCCATTGTTGTTCTTCCAGGGGAGCAGAGCGGAGCCGTCGTGGAGTCTG	60				
Db	1	ATGCTGTCCATTGTTGTTCTTCCAGGGGAGCAGAGCGGAGCCGTCGTGGAGTCTG	60				
QY	61	CTCCTTGAGGAGCCCTCCCGGACATGGAGGTGGAGTGAATCATCATCAGGACAA	120				
Db	61	CTCCTTGAGGAGCCCTCCCGGACATGGAGGTGGAGTGAATCATCATCAGGACAA	120				
QY	121	GGTCTGTCCAGTCTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC	180				
Db	121	GGTCTGTCCAGTCTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC	180				
QY	181	ATGCTTTATCCTAGGCGTGCACCTAAATGACACTGTCTTGTGGCACTCCTGGGGCCAAACC	240				
Db	181	ATGCTTTATCCTAGGCGTGCACCTAAATGACACTGTCTTGTGGCACTCCTGGGGCCAAACC	240				

Db 181 ATCTTTATCTAGGCTGCCACTAAATGGACACTCTCTGTGGCACTCTCTGGGGCAAAACC 240  
Qy 241 AAGCGCTGAGCTGTGGCCACCACTATCTGTGAACCTGTAGTGGCGGACCTGCTTTAT 300  
Db 241 AAGCGCTGAGCTGTGGCCACCACTATCTGTGAACCTGTAGTGGCGGACCTGCTTTAT 300  
Qy 301 GTGCTATTGGCCCTTCTCATCATCACTACCTACTAGATGATGAGTGGCCCTTGGGGAG 360  
Db 301 GTGCTATTGGCCCTTCTCATCATCACTACCTACTAGATGATGAGTGGCCCTTGGGGAG 360  
Qy 361 CTGCTCTGCAAGCTGGTGCACTCTCTGTCTATATCAACCTTTAGCGCAGCATCTGCTG 420  
Db 361 CTGCTCTGCAAGCTGGTGCACTCTCTGTCTATATCAACCTTTAGCGCAGCATCTGCTG 420  
Qy 421 CTGACCTGATCTCTGTGCAACAGTCTCTAGTGTGTGCCACCACTGTGTCGCTGCC 480  
Db 421 CTGACCTGATCTCTGTGCAACAGTCTCTAGTGTGTGCCACCACTGTGTCGCTGCC 480  
Qy 481 TACCGGACCCGAGCATGCTGTGGTGGGACACGACCACTGGCCCTGGTGGTCTC 540  
Db 481 TACCGGACCCGAGCATGCTGTGGTGGGACACGACCACTGGCCCTGGTGGTCTC 540  
Qy 541 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 541 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Qy 601 TATGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 660  
Db 601 TATGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 660  
Qy 661 TTGCTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720  
Db 661 TTGCTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720  
Qy 721 CCTGATCAAGCAGGAGGACCTCATGAGCAGGACGACGACGACGACGACGACGACGAC 780  
Db 720 CCTGATCAAGCAGGAGGACCTCATGAGCAGGACGACGACGACGACGACGACGACGAC 780  
Qy 781 CCGGACCACTCTACTGT 840  
Db 780 CCGGACCACTCTACTGT 840  
Qy 841 TCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
Db 840 TCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
Qy 901 AGCCAGTGTGCTCTACAGATATGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
Db 900 AGCCAGTGTGCTCTACAGATATGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
Qy 961 AGTCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAACTGA 1018  
Db 960 AGTCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAACTGA 1017

RESULT 2  
US-09-782-974C-13  
; Sequence 13, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogel, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1e1 G Protein Coupled Receptor  
; FILE REFERENCE: 411USPHRM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/165,838  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 60/198,568

; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 510  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-782-974C-13

Query Match 47.1%; Score 510; DB 9; Length 510;  
Best Local Similarity 100.0%; Pred. No. 2e-150;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGTAGTGGCGGACCTGCTTTATGTGCTA 306  
Db 1 TGGAGCTGTGCCACCACTATCTGTGAACCTGTAGTGGCGGACCTGCTTTATGTGCTA 60  
Qy 307 TTGGCCCTTCTCTATCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 366  
Db 61 TTGGCCCTTCTCTATCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120  
Qy 367 TGCAGCTGTGGTGCATCTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTGACC 426  
Db 121 TGCAGCTGTGGTGCATCTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTGACC 180  
Qy 427 TGCATCTGTGGCACCAGTTCCTAGTGTGTGCCACCACTGTGTGGCTGCCCTACCGG 486  
Db 181 TGCATCTGTGGCACCAGTTCCTAGTGTGTGCCACCACTGTGTGGCTGCCCTACCGG 240  
Qy 487 ACCCGCAGGATGCTGCTGGTGGGACCAAGCACCTTTGGGCTGGTGGTCCCTCCAGCTG 546  
Db 241 ACCCGCAGGATGCTGCTGGTGGGACCAAGCACCTTTGGGCTGGTGGTCCCTCCAGCTG 300  
Qy 547 CTGCCCACACTGGCTTCTCCCACACGACTACATCAATGGCCAGATGATCTGGTATGAC 606  
Db 301 CTGCCCACACTGGCTTCTCCCACACGACTACATCAATGGCCAGATGATCTGGTATGAC 360  
Qy 607 ATGACAGCCCAAGAGAATTTTGATGGCTTTTGCCCTACGCACTAGTTCTGACATTTGCT 656  
Db 361 ATGACAGCCCAAGAGAATTTTGATGGCTTTTGCCCTACGCACTAGTTCTGACATTTGCT 420  
Qy 667 GGCTTTCTTCCCTCTCTGCTGCTATTTTGGTGTGCTATTTACTGATGTCAGGACCTGAT 726  
Db 421 GGCTTTCTTCCCTCTCTGCTGCTATTTTGGTGTGCTATTTACTGATGTCAGGACCTGAT 480

RESULT 3  
US-09-791-932-32/c  
; Sequence 32, Application US/09791932  
; Publication No. US20030003451A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogel, Gabriel  
; APPLICANT: Parodi, Luis A.  
; APPLICANT: Hiebsch, Ronald R.







Db 500 CAACCTGGCCTTAGCCGATGGCTGGCCACACGACGCTGCTTCCAGAGTGCCAAAGTA 559  
Qy 333 ACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGACCTTCTGTCTTA 392  
Db 560 CCGTATGAGAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGACCTTCTGTCTTA 619  
Qy 393 TATCAACCTTTACGGCAGCAGCTGCTGCTGACCTGCTGCTGACCACTGCTTCTTAGG 452  
Db 620 CTACAATATGTTACACGAGCTTTCACGCTCACCATGATGAGTGTTCACCGCTACATCGC 679  
Qy 453 TGTGTGCCACCCACTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 512  
Db 680 TGTGTGCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739  
Qy 513 CAGCACCACCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563  
Db 740 CATCTGTATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790

RESULT 7

US-09-900-699A-1  
; Sequence 1, Application US/09900699A  
; Patent No. US20020162132A1  
; GENERAL INFORMATION:  
; APPLICANT: Brennan, Thomas J.  
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING DEZ ORPHAN  
; FILE REFERENCE: R-173  
; CURRENT APPLICATION NUMBER: US/09/900,699A  
; CURRENT FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/262,137  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/251,815  
; PRIOR FILING DATE: 2000-12-06  
; PRIOR APPLICATION NUMBER: US 60/219,403  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US 60/216,253  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1892  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-900-699A-1

Query Match 8.8%; Score 95.6; DB 9; Length 1892;  
Best Local Similarity 50.4%; Pred. No. 1.4e-19;  
Matches 233; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
Qy 154 GTCTACCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213  
Db 379 GTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438  
Qy 214 GTCTTGTGGCAGCTCCTGGGGCAACCAAGCGCTGGAGCTGTGCCACCACTATCTGGTG 273  
Db 439 GTGATTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498  
Qy 274 AACCTGATGCTGGCGGACCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333  
Db 499 CTGCTGTGGCGGACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558  
Qy 334 CTAGTACAGAGTGGCCCTTTCGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 393  
Db 559 ATGACTACCACTGGGTGCTGGGAAGCCAGTGCAGATGACCACTTCTTCTGCTCAGC 618  
Qy 394 ATCAACCTTTACGGCAGCAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453  
Db 619 CACAACATGTACACAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678  
Qy 454 GTGTGCCACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513

Db 679 GTGCTGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738  
Qy 514 AGCACCACTGGGCGCTGGTGGTCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573  
Db 739 TCGGCCCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798  
Qy 574 GACTTACATCAATGCGCCAGATGATCTGGTATGACATGACCAAGC 615  
Db 799 GCCAACATTGCGGAAGATAAAGCTGCTTCAACAACCTTCAGC 840

RESULT 8

US-09-779-679-23  
; Sequence 23, Application US/09779679  
; Publication No. US20030082757A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Mishra, Vishnu S  
; APPLICANT: Casman, Stacie  
; APPLICANT: Ballinger, Robert  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Li, Li  
; APPLICANT: Andrew, David P  
; APPLICANT: Mezes, Peter S  
; TITLE OF INVENTION: No. US20030082757A1e1. Proteins and Nucleic Acids Encoding the  
; FILE REFERENCE: 15966-661  
; CURRENT APPLICATION NUMBER: US/09/779,679  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: USSN 60/181045  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: USSN 60/183191  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: USSN 60/180,929  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: USSN 60/219758  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: USSN 60/181339  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: USSN 60/181344  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: USSN 60/221341  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: USSN 60/181392  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: USSN 60/219585  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: USSN 60/181157  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 422  
; TYPE: DNA  
; ORGANISM: Gallus gallus  
US-09-779-679-23

Query Match 8.8%; Score 95.2; DB 9; Length 422;  
Best Local Similarity 55.0%; Pred. No. 9.4e-20;  
Matches 187; Conservative 0; Mismatches 153; Indels 0; Gaps 0;  
Qy 169 GCCTACAGTATCATCTTTATCTAGGCTGCCACTAATGGCACTGCTTGTGGCACTCC 228  
Db 52 GTCTTACGATGTTGTTGCTCGGCTCATAGCACTGCTGCTATCTACATTTT 111  
Qy 229 TGGGGCCAAACCAAGCGCTGGAGCTGCGCACCACTATCTGGTGAACCTGATGGTGCC 288  
Db 112 ACTTTTACATTTGAAGTGGGAAGAGAGACGACGATACATGCTGAAATTTGGCGATATCG 171  
Qy 289 GACCTGCTTTATGTGCTATTGCTGCTTCTCTCATCATCACTACTACTAGATGACAGGTGG 348

Db	718	TGCTCTGTGCCACTGGGGCCCGTGGCTCCGCACAGCGCTGGCGCCGACGGCTGGTCTCGG	777
Qy	515	GCACCACTGGGCCCTGGTGGTCTCCAGCTGCTGCCACACTGGCCCTTCCTCCACACAGG	574
Db	778	TGGCCATCTGGACCTTGGTCTCTGCAGCAGATGCCCTTGCTCTTGATGGCCATGACCA	837
Qy	575	ACTACATCAATGCCCAGATGATCTGTATGTAGCATGACCAAGCAGAGAAATTTGATCGCC	634
Db	838	AGCGCTGGTGGGCAGCTGGCCCTGCATGGAGTACAGCAGCATGAGTCAGTCTCTCGGCC	897
Qy	635	TTTTTGGCCTAAGGCATAGTCTGTGACATTTCTCTGGGCTTCTCTTCCTCCCTTGG--TCATTTT	692
Db	898	TGCCCTCATGGTCTCGGTGGCCCTTTGGCCATTTGGCTTCTGTGGGCCAGTGGGGGATCATCC	957

```

Db      958  TGCTCTGCTATATGAAGATCACCTGGAAGCTGTGCAGCACAGCTGGGAGAACCCAGTGAC 1017
Qy      753  CAG 755
Db      1018  CAG 1020

RESULT 10
US-09-788-133-1
; Sequence 1, Application US/09788133
; Patent No. US20020052001A1
; GENERAL INFORMATION:
; APPLICANT: GLAXO GROUP LTD
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: P79011
; CURRENT APPLICATION NUMBER: US/09/788,133
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1020)
US-09-788-133-1
Query Match      8.2%; Score 88.4; DB 10; Length 1020;
Best Local Similarity 51.6%; Pred. No. 2e-17;

```

	malchines	220;	conservative	0;	flismacches	211;	indels	3;	gaps
Qy	161	TCCTCCTGGCCTACAGATATCATCTTTATCTTAGGGTGCACATAAATGGCACTGCTCTTGT	220						
Db	101	TGCGCTCCTTCTACCTTCTGGATTTTATCTTGGCTTAGTTGGCAATACCTCGGCTCTGT	160						
Qy	221	GGCATCTCTGGGGCCCAACCAAGCGTGGAGCTGTGCCACCACTTATCTGTGTGAACCTGA	280						
Db	161	GGCTTTTCATCCGAGACCACAAGTCCGGGACCCCGGCCAACGTGTTCTGATGCATCTGG	220						
Qy	281	TGCTGGCCGACCTGCCTTTATGTGCTAT----TGCCCTTCTCATCAATCAACCTACTCACATAG	337						
Db	221	CCGTGGCGCACTTGTCTGGTGGTCTGTGCTTGCCTGCCACCGCGCTGTGCTTACCACCTTCTCTG	280						
Qy	338	ATGACAGGTGGCCCTTCGGGGAGCTCTCTGCAAGCTGGTGGACATTCCTCTGTTCTATATCA	397						
Db	281	GGAAACCATGGCCATTTGGGGAAATCGCATGGCTCTCACCGCGTTCCTCTTCTACCTCA	340						
Qy	398	ACCTTTACGGCAGCATCCTGCTGCTGACCTGCATCTCTGTGCACAGTTCCTAGGTGTGT	457						
Db	341	ACATGTACGGCAGCATCTACTTCTCCACCTGCATCAGCGCGCACCGTTCCTTGGCCATTTG	400						
Qy	458	GCACACCACCTGTGTTTCGTGCGCCCTACCGCACCGCAGAGCATGCCGTGGCCACCA	517						
Db	401	TGCACCGGTCAAGTCCCTCAAGCTCCGACGGCCCTCTACGCACACCTGGCCGTGGCCT	460						
Qy	518	CCACCTGGGGCCCTGGTGGTCTCCAGCTGCTGCCACACATGGCCCTTCTCCACACAGGACT	577						

Qy	265	TATCTGTGAACATGATGTTGGCGACCTCCTTATGTGCTATATGCCCTTCCTCATCATC	324
Db	633	TTCCTGTTCACCATGGCCGTGGCGACCTCCTGC---TGGCTTCATCTTGGCCTTTTGC	689
Qy	325	ACCTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC	384
Db	690	GTGCCCAGGGCTCTGTGGGCTGGTCTGGGACCTTCCTCTGCAAACTGTGATTGCTGC	749

Qy	385	CTGTTCTATATCAACCTTTACGCAGCATCTCTGCTGACCTGCACTCTCTGTGCACCAAG	444
Db	750	CTGCACAAAGTCAACTTCTACTGCAGCAGCGCTCTCTGSCCTGCATCGCGTGGAGCCGC	809
Qy	445	TTCTTAGGTGTGGCCACCACACTGTGTTCCGTGCCCTACCGGACCCGCGAGGCATGCGCTGG	504
Db	810	TACCTGGCCATTGTCCACGCGCTCCATGCGCTACCGCACCGCGCGCTCTCTCCATCCAC	869
Qy	505	CTGGGCACCAAGCAACCTGGGCGCTGGTGGTCTCCAGCTGCTGCCACACACTGGCGCTTC	564
Db	870	ATCACTGTGGGACCATCTGGCTGGTGGGGCTTCTCTTGCCCTTGCCAGAGATTCTCTTC	929
Qy	565	TTCCA	569
Db	930	GCCAA	934

```

RESULT 12
US-09-728-422-1
: Sequence 1, Application US/09738422
: Patent No. US20020128187A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Ren, Feiyan
: APPLICANT: Zhao, Qing A.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wehrman, Tom
: APPLICANT: Drmanac, Radojé T.
: TITLE OF INVENTION: No. US20020128187A1el Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 787CIP2F
: CURRENT APPLICATION NUMBER: US/09/728,422
: CURRENT FILING DATE: 2000-11-30
: PRIOR APPLICATION NUMBER: 09/560,875
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 09/496,914
: PRIOR FILING DATE: 2000-02-03
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: pt_FL_genes Version 2.0
: SEQ ID NO 1
: LENGTH: 1313
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (547)..(1239)
US-09-728-422-1

```

	Query Match	7.9%;	Score 86;	DB 10;	Length 1313;
	Best Local Similarity	46.7%;	Pred. No. 1.3e-16;		
	Matches 309;	Conservative	0;	Mismatches 350;	Indels 3;
	Gaps 1;				
Qy	147	CAAGCAAGTCTACCTCTCCCTGGCCTACAGATCATCTTTATCTCTAGGGCTGCCACCTAAA	206		
Db	639	CAAGATGCACATACCTCCCTGTTATTTATGGCATTTATCTCTCTGGGATTTCCAGGCCAA	698		
Qy	207	TGGCAGCTGCTTGTGGGACCTCCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACTA	266		
Db	599	TGCAGTAGTGATATCCAGCTTACATTTTCAAATGAGACCTTTGGAAGAGACGACACCATCAT	758		
Qy	267	TCTGGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTAT---TGCCCTTCTCTCATCAT	323		

Db 759 TATGCTGAACCTGGCCCTGCACAGATCTGCTATCTGACAGCCTCCCTTCTGATTC 818  
Qy 324 CACCTACTACTAGATGAGCTGCGCCCTTCGGGAGCTGCTGCAAGCTGGTGCATTT 383  
Db 819 CTATATGCCAGTGGGAAACTGATCTTTGGAGATTTCAATGTAAGTTATCCGCTT 878  
Qy 384 CCTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCAFTCTGTGCACCA 443  
Db 879 CAGCTTCCATTTCAACCTGTATAGCAGCATCTCTTCTCCTCACCTGTTTCAGCATCTTCCG 938  
Qy 444 GTTCTAGGTGTGCCACCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503  
Db 939 CTACTGTGTGATCATTTACCAATGAGTGTCTTTTCCATTTCACAAACTCGATGTGCA 998  
Qy 504 GCTGGGACCAAGCACCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 563  
Db 999 TGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058  
Qy 564 CTCACACAGGACATCAATGACAGATGATGCTGATGATGATGATGATGATGATGATGAT 623  
Db 1059 GATCATATCAACCAAGGACCAAGATGATGATGATGATGATGATGATGATGATGATG 1118  
Qy 624 TTTGATCGGCTTTTGGCTAGCGATAGTTCTGACATTTCTGGCTTTCTTCCCTCCT 683  
Db 1119 ACTCAATACTAATGAGTGTGATCAACCTAAATTTTGACTGCAACTACTTTGCTCCCTTGG 1178  
Qy 684 TGGTCAATTTTGGTGTGCTATTCACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743  
Db 1179 TGATAGTACACTTTGCTATACCAAGATATCCACACTGCTGACCCATGCTGCAACTG 1238  
Qy 744 TCATGAGACAGGACCAACACAGCCGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 803  
Db 1239 ACAGTGTCTTAAAGCAGAAAGCAGGAGGCTTAACCATTTCTGCTACTCCCGACGCTGG 1298  
Qy 804 GC 805  
Db 1299 TC 1300

RESULT 13  
US-09-864-761-21546/c  
; Sequence 21546, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aesomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 21546  
; LENGTH: 348  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL009181.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2  
; OTHER INFORMATION: EST\_HUMAN HIT: BE252309.1, EVALUE 2.10e-02  
; OTHER INFORMATION: NT HIT: g11526121, EVALUE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: P41143, EVALUE 6.00e-48  
US-09-864-761-21546

Query Match 7.9%; Score 85.8; DB 10; Length 348;  
Best Local Similarity 55.6%; Pred. No. 7.8e-17;  
Matches 165; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
Qy 236 AACCAAGCGCTGGAGCTGTGCCACCACTATCTGTTGTAACCTGTGTCGCCGACCTGC 295  
Db 301 ACATAAGATGAAGAGCGGCCACCAACATCTACATCTTCAACCTGGCTTAGCCGATGCC 242  
Qy 296 TTTATGCTATTGCCCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCG 355  
Db 241 TGGCCACCAAGCGCTGTGCCAGAGTGCCAAAGTACCTGATGAGAGCTGGCCCTTCG 182  
Qy 356 GGGAGCTGCTGCAAGCTGTGCACCTTCCCTGTTCTATATCAACCTTTACGGCAGCATCC 415  
Db 181 CGAGCTGCTGCAAGCTGTGCTCTCCATCGACTACTACAATATGTTACACGACATCT 122  
Qy 416 TGCTGCTGACCTGCTGCTGTGCCAGGATTTCTAGTGTGTGTCACCACTGTGTTCG 475  
Db 121 TCACGCTCACCATGATGATGTTTTCACCGCTACATCGCTGTCTGCCACCTCTCAAGGCC 62  
Qy 476 TGGCCCTACCGAGCCGAGGATGCTGCTGGTGGCCAGCAGCAGCAGCAGCAGCAGCAG 532  
Db 61 TGGACTTCCGACGCTGCCAAGCCCAAGCTGATCAACATCTGTATCTGGTCTCTGG 5  
RESULT 14  
US-09-823-114-7  
; Sequence 7, Application US/09823114  
; Patent No. US20020061554A1  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, CHRISTOPHER J.  
; APPLICANT: KEITH, DUANE E.  
; TITLE OF INVENTION: OPIOID RECEPTOR GENES

POPULOGR: Linear	MOLECULE TYPE: DNA (genomic)	FEATURE:	NAME/KEY: CDS	LOCATION: 58..1173	US-09-214-904-3
Query Match	7.7%	Score 83.6;	DB 10;	Length 2218;	
Best Local Similarity	47.7%	Pred. No. 9.1e-16;			
Matches 275;	Conservative 0;	Mismatches 299;	Indels 2;	Gaps	
QY	155	TCTACCTCTCCTCGCCCTACAGTATCATCTTTATCTCTAGGGCTGCCACTAAATGGCACTG	214		
Db	206	TGCGCATCACCGCGCTCTACTTCGGCTGTGTGCGCAGTGGGGCTTC	265		
QY	215	TCATTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACTATCTGGTGA	274		
Db	266	TCATGTTTGGGATCGTCCGGTACACAAATGGAAGACCGCCCAACATCTACATCTTCA	325		
QY	275	ACCTGATGGTGGCCGACCTGCTTTATGTGCTATTGGCCCTTCCTCATCATCACCTACTCAC	334		
Db	326	ATCTGGCTTTGGCTGATGCGCTGGCCACCAAGCAGCTGCCCTTCCAGAGCGCAAGTACT	385		
QY	335	TAGATGACAGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTCCTCTGTCTATA	394		
Db	386	TCATGGAAACGTGCCCGCTTTGGCGAGCTGCTGTGCAAGGCTGTCTCTCCATTGACTACT	445		
QY	395	TCAACCTTTACGGCAGACATCTGCTGCTGACCTTCATCTCTGTGGCAACCACTTCCTAGGTG	454		
Db	446	ACAACATGTTCACTAGACATCTTACCCTCACCATGATGACCGTGGACCGCTACATTGCTG	505		
QY	455	TGTGCCACCCCACTGTGTGCTGCGCTACCGGACCCCGACGGCATGCTGGCTGGGCACCA	514		
Db	506	TCTGCACTCTGTCAAAAGCCCTTGGACTTCGGACACCAAGCAAGGCCAAGCTGATCAATA	565		

Qy	515	GCACCACCTGGGGCCCTGGTGGTGCTCCAGCTGCTGCCACACTGGCCCTTCCTCCACACGG	574
Db	566	TATGCATCTGGGTCTTTGGTTCAGGTGTCGGGGTCCCATCATGTCATGGCAGTGACCC	625
Qy	575	ACTACATCAATGGCCAGATGATCTGGTATGACATGACACGCAAGAGAAATTTTGATCGCC	634
Db	626	AACCCGGGATGGTGCAGTGGTATGCATGCTCCATTTCCCACTCCCACTGGTACTGGG	685
Qy	635	TTTTTGCCTACGGCATAGTCTTGACATTTGCTGGGCTT--TCTTTCCCTCCTTGGTCAATT	692
Db	686	ACACTGTGACCAAGATCTCGTGTTCCCTTTTGGCCTTCTGGTGGCGATCCTCATCATCA	745
Qy	693	TGGTGTGCTATTTCATGTATGGTCAGAGCCTGATCA	728
Db	746	CGGTGTGCTATTGGCCTCATGCTACTCGCCTCGCA	781

Search completed: May 26, 2003, 22:57:54  
Job time : 213 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 21:20:18 ; Search time 1815 Seconds  
(without alignments)  
9654.830 Million cell updates/sec

Title: US-10-023-586b-1  
Perfect score: 1082  
Sequence: 1 atgtgttcatttgccttcc.....ccagggttgacagatctgg 1082

oring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_nam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	200.6	18.5	2542	11 AK017378	AK017378 Mus muscu
2	200.6	18.5	3001	11 AK005013	AK005013 Mus muscu
3	141.2	13.0	641	14 BQ396255	BQ396255 NISC_ngl9
4	140.2	13.0	638	9 AL675845	AL675845 AL675845
5	128.2	11.8	899	14 BQ959110	BQ959110 AGENCOURT
6	123.2	11.4	623	14 BQ038875	BQ038875 pgnic.pk0

7	121.6	11.2	851	13 BI833118	BI833118 603090834
8	120.6	11.1	928	9 AL561864	AL561864 AL561864
9	120.4	11.1	613	9 AL657842	AL657842 AL657842
10	120.2	11.1	1063	14 BM918491	BM918491 AGENCOURT
11	119.8	11.1	408	13 BI401676	BI401676 MI-P-CPD-
12	119.6	11.1	1101	17 CNS0532S	AL18925 Tetraodon
13	113.8	10.5	955	9 AL521440	AL521440 AL521440
14	109.8	10.1	898	13 BI819396	BI819396 603034571
15	106.2	9.8	884	9 AL525099	AL525099 AL525099
16	106	9.8	553	13 BM031311	BM031311 496677 MA
17	103	9.5	606	17 AZ953874	AZ953874 2M02191L7
18	102.8	9.5	934	13 BI768397	BI768397 603053844
19	99.4	9.2	932	17 CNS0251R	AL211932 Tetraodon
20	97.8	9.0	484	14 BQ551383	BQ551383 H4008H11-
21	97.6	9.0	936	9 AL520218	AL520218 AL520218
22	96.6	8.9	1118	14 BM918898	BM918898 AGENCOURT
23	94.6	8.7	835	9 AL525098	AL525098 AL525098
24	94.2	8.7	663	13 BM426517	BM426517 pgf2n.pk0
25	91.6	8.5	467	10 AW323925	AW323925 ug41d11.y
26	91.4	8.4	720	13 BI753905	BI753905 603022907
27	90.4	8.4	491	12 BG712193	BG712193 pg11n.pk0
28	89.4	8.3	801	9 AL549441	AL549441 AL549441
29	89.2	8.2	549	12 BE938149	BE938149 MR3-TN004
30	89	8.2	764	12 BF345996	BF345996 602018717
31	88.8	8.2	609	9 AL588350	AL588350 AL588350
32	88.8	8.2	801	13 BG924078	BG924078 602823635
33	88.6	8.2	695	9 AJ450320	AJ450320 AJ450320
34	88.6	8.2	719	9 AJ455780	AJ455780 AJ455780
35	88.4	8.2	618	12 BF346620	BF346620 602021219
36	88.4	8.2	663	12 BG820137	BG820137 602782353
37	88.4	8.2	691	13 BG912683	BG912683 602807060
38	88.4	8.2	779	12 BG820495	BG820495 602782813
39	88.4	8.2	788	13 BI916752	BI916752 603178755
40	88.4	8.2	1083	14 BM805382	BM805382 AGENCOURT
41	88.4	8.2	1136	14 BM906977	BM906977 AGENCOURT
42	87.4	8.1	961	14 BQ056144	BQ056144 AGENCOURT
43	87.4	8.1	1139	14 BM922349	BM922349 AGENCOURT
44	86.8	8.0	678	10 BE314834	BE314834 601147561
45	86.2	8.0	1064	13 BI517798	BI517798 603042051

ALIGNMENTS

RESULT 1	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
LOCUS	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
DEFINITION	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
ACCESSION	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
VERSION	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
KEYWORDS	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
SOURCE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
ORGANISM	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
REFERENCE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
AUTHORS	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
TITLE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
JOURNAL	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
MEDLINE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
PUBMED	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
REFERENCE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
AUTHORS	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
TITLE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
JOURNAL	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
MEDLINE	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002
PUBMED	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	HTC 19-JAN-2002



```

QY 604 GACATGACCAGCAAGAGAAATTTTGATCGCTTTTTCCTACGGCATATGTTCTG--ACAT 661
|||||  || ||||| |||||  || ||||| || ||||| || |||||
Db 1036 GACACCTCGCCGCGAGAGCTCTTTAGCCATTTTGTGGCTTACAGCTCCGCTCATGCTGGGT 1095

QY 662 TGTCTGGCTTCTTCTCCCTCTGGTTCATTTGGTGTGCTATTACATGATGGTCAGAGC 721
|||||  || ||||| || ||||| || ||||| || ||||| || |||||
Db 1096 CTGCTTTTGTCTGGCTTCTCCGTAATCTGCTGTGTACGTGTATGGCCAGCGCG 1155

QY 722 CTGATCAAGCCAGGAGAGAACCTTCATGAGGACAGGCAACACGCCGAGCGTCCATC 781
|||||  || ||||| || ||||| || ||||| || ||||| || |||||
Db 1156 CTGCTCAACCGGCTTATGGACACAGGAGGTCTGCCCTGGGCCAAGCCAGTCTGTG 1215

QY 782 CGGACCATCTACTGTGTGTGGCTTCTACCCCTCTGTGTTGTGCGCTTCCATATCACT 841
|||||  || ||||| || ||||| || ||||| || ||||| || |||||
Db 1216 CGCACCATCTGCTGTGTGTGCTGCTCTCGCCCTCTGCTTCTGCTTTCACCTCACG 1275

842 CGCTCTCTTACCTACCATCTGCTCTCTCTGCTTCTCAGGACTGCCAGCTTCTGATGGCA 901
|||||  || ||||| || ||||| || ||||| || ||||| || |||||
Db 1276 CGCACCTCTCTACTCTCTCTCTCTCGAT--CACTTGACCTCAGCTGCGCACACCTCAACGCC 1332

QY 902 GCCAGTGTGGCTTACAGATATGAGGCGCTCTGCTGAGTGTGACGAGCTGCCCTCAACCCCA 961
|||||  || ||||| || ||||| || ||||| || ||||| || |||||
Db 1333 ATCAACATGSCATATAAGATCACCGCGCGCTGCGCAGCGCCCAACAGTTGCTTGACCCG 1392

QY 962 GTCTGTACTTCTTCTCAAGG 982
|||||  || ||||| || ||||| || ||||| || ||||| || |||||
Db 1393 GTACTCTACTTCTGCGCAGG 1413

RESULT 2
AK005013 3001 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male liver cDNA, RIKEN full-length enriched
DEFINITION library, clone:1300015C04:purinergic receptor P2Y, G-protein
coupled 2, full insert sequence.
AK005013
VERSION AK005013.1 GI:12836638
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA,
clone:1300015C04.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED 10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

```

# TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gofobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalow, S., Casavant, T., Fleschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, O., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringuwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5 (bases 1 to 3001)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

# TITLE JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

# COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGCGCCGCACTCGAGTTTTTTTTTTTTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGAGAGAGAGGATCAAGAGCTCAATTAATTAATTAACCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOUR.

# FEATURES

Location/Qualifiers  
1. .3001  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:1300015C04"  
/db\_xref="MGI:1895057"  
/db\_xref="taxon:10090"  
/clone="1300015C04"  
/sex="male"  
/tissue\_type="liver"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1. .3001  
/gene="P2ry2"

# gene

```
CDS
477..1598
/ gene="P2ry2"
/ note="data source:MGD, source key:MG1:105107,
evidence:ISS
purinergic receptor P2Y, G-protein coupled 2
putative"
/codon_start=1
/protein_id="BAB23746.1"
/db_xref="GI:12836639"
/translation="MAADLEPNSTINGTNEGDELGYKCFNEDFKVLLPVSGVVC
TVLGLNVALYITFLCKATKNASTYMFHLVAVSDSYAASLELLYIYARGHWPFS
VLAKLVFLYFTVNRGTRITRCHDSARELSHFVAYSSVGLGLFPAVFPVILVCY
VLMARLLPAYGTGGLPRKRSVRTIALVLAVFALCFLPFHLTRTLYSFRSLDL
SHTLAINMAYKITRPLANSCLDPVLYFLAGORLVRFARDAKPPTPTSPQARR
KLGHRPNRTVRKDLVSSDDSRTESTPAGSEKDIRL"
polyA_signal
2982..2987
/ gene="P2ry2"
/ note="putative"
3001
polyA_site
/ gene="P2ry2"
/ note="putative"
3011
BASE COUNT      630 a      861 c      834 g      676 t
IN
Query Match      18.5%; Score 200.6; DB 11; Length 3001;
Best Local Similarity 54.5%; Pred. No. 1.2e-43;
Matches 469; Conservative 0; Mismatches 384; Indels 8; Gaps 3;
Qy 127 TGCAGATTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCGCTACAGTATCATCTTT 186
Db 549 TGTGCTTTCAACAGAGACTTCAAGTAGCTGTGTGGCCGTGTCTATGGGTGTGTGC 608
Qy 187 ATCTAGGCTGCCACCAATGACCTGTCTGTGGCACTCTCTGGGCCAAACCAAGCGC 246
Db 609 GTGCTGGGTGTGCTGAAGCTGTGGCTCTCTATATCTTCTATGCGGCTCAAAACC 668
Qy 247 TGGAGCTGTGCCACCACTATCTGTGTAACCTGATGTGGCCGACCTGCTTTATGTGCTA 306
Db 669 TGGAGCGCTCCACCACTATCTGTTCACCTGGCAGTTTCCGAGTCTCTCTACGACGC 728
Qy 307 T---TGGCTTCTTCATCATCACTACTACTACTAGATGACAGTGGCCCTTGGGAGCTG 363
Db 729 TCCCTGGCGCTGTGGTTTATCTACGCCGGGTGACCACTGCGCATTTAGCAGCGTG 788
Qy 364 CTCTGCAAGCTGGTGCATCTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTGCTG 423
Db 789 CTCTGCAAGCTGGTGGTTCCTCTCTTACACCAACCTCTACTGACGATCCTCTCTC 848
Qy 424 ACCTGCACTCTGTGCACCACTCTCTAGGTGTGTGCCACCACTGTGTGCTGCGCTAC 483
Db 849 ACCTGCACTGAGCTGCACCGGTGCTGGAGTCTCTGGCCCTCTGCACTCCCTGCGTTG 908
Qy 484 CGGACCCGACGATGCTGCTGGCAGCAGCAGCACCCTGGCCCTGTGGTCTCTCCAG 543
Db 909 GCGCGGCCGCTTATGCGCGCGGTGGCTGCGGTGTGTGGGTGCTGGTGTGCGCTGCG 968
Qy 544 CTGCTGCCCACTGCGCTCTCCACACGCACTACATCAATAGCCAGATGATCTGCTAT 603
Db 969 CAGGACCCGCTCTACTCTCTGTCACCAACGAGCTGGCGGGAACCCGGATCCTTGCCAT 1028
Qy 604 GACATGACCAAGCAAGAAATTTTATGCGCTTTTGGCTACGGCAGATAGTTCTG--ACAT 661
Db 1029 GACACTGCGGCCGAGAGCTCTTTAGCCATTTTGTGGCTTACAGCTCGCTATGCTGGGT 1088
Qy 662 TGTCTGCTTTCTTCTCTCTGCTATTTTGTGTGCTATTCTACATCATGCTGAGGAGC 721
Db 1089 CTGCTTTTGTGTGCGCTTTTTCGTAATCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1148
Qy 722 CTGATCAAGCCAGAGGAAACCTCTATGAGGACAGCAACAGCCGAGCCAGGCTCCATC 781
Db 1149 CTGCTCAACCGCGCTTATGGGACACACAGGAGGTCTGCTCGGGCCAGCGCAAGTCTGTG 1208

RESULT 3
BQ396255
LOCUS      NISC_ng19g07.y1 NICHD XGC Emb6 Silurana tropicalis cDNA clone
DEFINITION IMAGE:5383884 5', mRNA sequence.
ACCESSION  BQ396255
VERSION     BQ396255.1 GI:21083932
SOURCE      western clawed frog.
ORGANISM    Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodidae; Silurana.
REFERENCE   1 (bases 1 to 641)
AUTHORS    NIH-XGC http://image.llnl.gov/image/html/xenopuslib_info.shtml.
TITLE      National Institute of Child Health and Human Development, National
           Cancer Institute, Xenopus Gene Collection
JOURNAL     Unpublished (2002)
COMMENT     Contact: Robert Strausberg, Ph.D.
           Email: cgaps-remail.nih.gov
           cDNA Library Preparation:
           NIH-XGC http://image.llnl.gov/image/html/xenopuslib_info.shtml
           DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL
           Sequencing Center (NISC)
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           info@image.llnl.gov
           Plate: L1AM11977 row: N column: 13
           Seq primer: M13RPI reverse primer (ABI).
FEATURES
     source
     1..641
     /organism="Silurana tropicalis"
     /db_xref="taxon:8364"
     /clone="IMAGE:5383884"
     /clone_lib="NICHD XGC Emb6"
     /tissue_type="neural"
     /dev_stage="embryo, stages 14-19"
     /lab_host="DH10B (phage-resistant)"
     /notes="vector: pCMV-SPORT6.cdb; Site_1: NotI; Site_2:
     EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average
     insert size 2.1 kb. Constructed by Invitrogen. Note: This
     is a Xenopus Gene Collection (XGC) library."
BASE COUNT      132 a      184 c      153 g      172 t
ORIGIN
Query Match      13.0%; Score 141.2; DB 14; Length 641;
Best Local Similarity 55.5%; Pred. No. 1.2e-27;
Matches 294; Conservative 0; Mismatches 233; Indels 3; Gaps 1;
Qy 127 TGCAGATTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCGCTACAGTATCATCTTT 186
Db 106 TGTAAATTCGACAGAGATTTTCAAGTAGCTCTCTCCCGCTGTGTACGGCATCGTGTC 165
Qy 187 ATCTAGGCTGCCACATAAATGGCACTGTCTGTGGCACTCTCTGGGCGCAACCAAGCGC 246
```

Db 166 TCGGTGGGCTGATCTGAACATFTGGGCCCTGTACATCTTCTCTGTTCGGATCAAGCCC 225  
QY 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTA 306  
Db 226 TGGAGCCCTCCACCACTACATGTTCAACCTGGCCATTTCCGACATGATGTACGTCATC 285  
QY 307 T---TGGCCCTTCCTCATCATACCTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTG 363  
Db 286 TCCTCCCGCTGCTGTCTATTACTACTCGCAGGGGACAACTGGCCGTTCGGCGCTGGCC 345  
QY 364 CTCTGCAAGCTGTGCACATTCCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTG 423  
Db 346 TTGTGCAAGATGCTCAAGTTCCTTTCTACACCAACATGTACTGTAGCATCTCTCTCTTG 405  
QY 424 ACCTGCATCTCTGTGCACAGTTCCTAGTGTGTGCACACCACTGTGTCTGCTGCGCTTAC 483  
406 CTGTGATCAGCATCACCGCTTCCTCGGATCTGCTACCCATGAATCACTCGGTTGG 465  
QY 484 CGGACCCGAGGATGCTGTGGTGGGACACGACACCTGGGCCCTGTGGTTCCTCCAG 543  
Db 466 CTGAAGGTGCGGAATGCTCGGATTAATTCGTGTGCTGCTGGTATTCGTTTCCGCTGT 525  
QY 544 CTGCTGCCACATGCTCCACAGGACTACATCAATGCCAGATGATCTCGTAT 603  
Db 526 CAGTCCCTATATTGTACTTTGTACACACAGTTCCCAATGGGACAGCACCTGCCAT 585  
QY 604 GACATGACCAAGCAAGAAATTTGATCGCTTTTTCGCTACGGCATAGT 653  
Db 586 GACACTTCCAGCGTGATCTATTGTGACAACTTTGTGCTGTACAGACAGT 635

RESULT 4  
AL675845  
LOCUS  
DEFINITION AL675845 XGC-gastrula Silurana tropicalis cDNA clone Tgas05ln19 5',  
mRNA sequence.  
ACCESSION AL675845  
VERSION AL675845.1 GI:19532219  
KEYWORDS EST.  
SOURCE western clawed frog.  
ORGANISM Silurana tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
Xenopodinae; Silurana.  
1 (bases 1 to 638)  
Taylor R., Ashurst J.L., Croning M.D.R., Zorn A.M. and Rogers, J.  
Sanger xenopus tropicalis EST project 2002  
Unpublished (2001)  
Contact: Taylor R  
Sanger Centre  
Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: trop@sanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS\_SEQUENCE\_ID: Tgas05ln19.picSP6  
Sequencing primer: PICSP6  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn.  
Location/Qualifiers  
1. 638  
/organism="Silurana tropicalis"  
/db\_xref="taxon:8364"  
/clone="Tgas05ln19"  
/clone\_lib="XGC-gastrula"  
/dev\_stage="gastrula (stages 10 5-13 mixed)"  
/lab\_host="Escherichia coli XL1-blue"  
/note="Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA  
was oligo dT primed from 5ug of poly A+ RNA from stages  
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated  
into pCS107 with EcoRI at the 5' end and NotI at the 3'  
end."

BASE COUNT 134 a 180 c 144 g 179 t 1 others  
ORIGIN

Query Match 13.0%; Score 140.2; DB 9; Length 638;  
Best Local Similarity 55.3%; Pred. No. 2.3e-27;  
Matches 293; Conservative 0; Mismatches 234; Indels 3; Gaps 1;  
QY 127 TGGCAGTTCTCAGAGATACAGCAAGTCTACTCTCCCTGGCCTACATATCATCTTT 186  
Db 102 TGTAAATTCAGCAGGATTTCAAGTACGTCTCTCCCGTGTGTCAGGCATCGTGTTC 161  
QY 187 ATCTTAGGCTGCCATTAATGGCACTGTCTGTGGCACTCCCTGGGCGCAACCAAGCGC 246  
Db 162 TGGCTGGGCTGATCTTGAACATCTTGGCCCTGTACATCTTCTGTTCGGATCAAGCCC 221  
QY 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTA 306  
Db 222 TGGAGCCCTCCACCACTACATGTTCAACCTGGCCATTTCCGACATGATGTACGTCATC 281  
QY 307 T---TGGCCCTTCCTCATCATCACCTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTG 363  
Db 282 TCCTCCCGCTGCTGTCTATTACTTCGACGGGGACAACCTGGCGTTCGGCGTGGCC 341  
QY 364 CTCTGCAAGCTGTGCACATTCCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTG 423  
Db 342 TTGTGCAAGATGCTCAAGTTCCTTTCTACCAACATGTACTGTACATCTCTCTCTTG 401  
QY 424 ACCTGCATCTCTGTGCACCACTTCCTAGTGTGTGCACACCACTGTGTTCGCTGCCCTAC 483  
Db 402 CTGTGATCAGCATCCACCGCTTCCTCGGATCTGCTACCCAATGAATCACTCGGTTGG 461  
QY 484 CGGACCCGAGGATGCTGTGGTGGGACACGACACCACTGGGCCCTGTGGTTCCTCCAG 543  
Db 462 CTGAAGGTGCGGAATGCTCGGATTAATTCGTGTGCTGCTCGGTTATCTGTTTCGGCTGT 521  
QY 544 CTGCTGCCACATGCTGCTTCCTCCACAGGACTACATCAATGCCAGATGATCTGCTAT 603  
Db 522 CAGTCCCTATATTGTACTTTGTACACCACTTTCCAGTTCGGGACACCACTGCCAT 581  
QY 604 GACATGACCAAGCAAGAAATTTGATCGGCTTTTTCGCTACGGCATAGT 653  
Db 582 GACACTTCCAGCGTGGATCTATTGTGACAACTNTGTGTCTACAGCACAGT 631

RESULT 5  
BO959110  
LOCUS  
DEFINITION BO959110 899 bp mRNA linear EST 21-AUG-2002  
5', mRNA sequence.  
ACCESSION BO959110  
VERSION BO959110.1 GI:22374588  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 899)  
NIH-MGC http://mgc.ncl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: LLCM2651 row: e column: 10  
High quality sequence stop: 677.  
Location/Qualifiers  
1. .899  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6462033"

/clone\_lib="NIH\_MGC\_101"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pOTB7; Site:1: EcoRI; Site\_2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."  
 BASE COUNT 131 a 292 c 301 g 175 t

Query Match 11.8%; Score 128.2; DB 14; Length 899;  
 Best Local Similarity 60.08; Pred. No. 5-le-24;  
 Matches 250; Conservative 0; Mismatches 163; Indels 4; Gaps 2;  
 Qy 127 TGCAGTTCTCAGAGAGTACAAGAGTACCTCTCCCTGGCTACAGTATCATCTTT 186  
 Db 405 TCCCGCTTCAACGAGGACTTCAAGTACGTGCTGCTGTCTCTACAGCGTGTGTGC 464  
 Qy 187 ATCTAGGCTGCCACCACTATCTGTGGCACTCTCTGTGGCACTCTCTGGGCAACCAAGGCG 246  
 Db 465 GTGCTGGGCTGTGTCTGAACCGCGTGGCGCTCTACATCTTCTTGGGCGCGCTCAAGACC 524  
 Qy 247 TGGAGCTGTGCCACCACTATCTGTGTAACCTGATGTGGCGGCACTCTCTTTATGTGCTA 306  
 Db 525 TGAATGGGTCCACACATATATGTTCCACCTGGCTGTCTGATGCATCTATGCGGCC 584  
 Qy 307 T---TGCCCTTCTTCATCATCTACTACTACATAGATGACAGGTGGCCCTTCGGGAGCTG 363  
 Db 585 TCCCTGGCGCTGGTGTATTACTACGCCCGCGGCACTGGCCCTTCAGCACGGTG 644  
 Qy 364 CTCTGCAAGCTGGTGCACTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGTGCTG 423  
 Db 645 CTCGCAAGCTGGTGCGCTCTCTCTACACCAACCTTTACTGCGAGTCTCTTCTC 704  
 Qy 424 ACCGTCATCTCTGTGCA--CCAGTTCCTAGTGTGTGGCCACCACTGTGTCGCTGCCCTA 482  
 Db 705 ACCGTCATCAGCTGCACCCGGTGTCTGGCGGTCTTACGACCTCTGCGCTCCCTGCGCTG 764  
 Qy 483 CCGACCCGCGAGGATGCTGCTGGTGGGCACTGACGACCACTGGGCGCTGGTCTCT 539  
 Db 765 GGGCGGCGCGCTAGCTGCCGGGGGGCGGGGGCGGGGGGGGGTGTCT 821

RESULT 6  
 BQ038875 623 bp mRNA linear EST 01-MAY-2002  
 LOCUS  
 DEFINITION pgnlc.pk010.113 normalized chicken lymphoid cDNA library  
 gallus cDNA clone pgnlc.pk010.113 5' similar to  
 spIP34996[P2YR.CHICK P2Y PURINORECEPTOR 1 (ATP RECEPTOR) (P2Y1)  
 (PURINERGIC RECEPTOR) pIRIS3733 G protein-coupled receptor -  
 chicken embi[CA51716.1] (X73268) ATP receptor P2Y1 [Gallus gallus],  
 mRNA sequence.

ACCESSION BQ038875  
 VERSION BQ038875.2 GI:20383637  
 KEYWORDS EST.  
 SOURCE chicken.

ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 623)  
 AUTHORS Morgan, R.W. and Burnside, J.  
 TITLE Chicken lymphoid ESTs  
 JOURNAL Unpublished (2001)

COMMENT On Mar 27, 2002 this sequence version replaced gi:19772415.  
 Contact: Joan Burnside  
 Molecular Endocrinology  
 University of Delaware  
 40 Townsend Hall, Newark, DE 19717, USA

Tel: 302 831-1345  
 Fax: 302-831-3411  
 Email: joan@udel.edu, www.chickest.udel.edu.  
 FEATURES  
 source Location/Qualifiers  
 1..623

/organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /clone\_lib="pgnlc.pk010.113"  
 /sex="Male and Female"  
 /tissue\_type="thymus, bursa, spleen, PBL, bone marrow"  
 /lab\_host="E.coli EMDH10B"  
 /note="vector: pCMVSPORT 6"  
 BASE COUNT 117 a 199 c 172 g 128 t

Query Match 11.4%; Score 123.2; DB 14; Length 623;  
 Best Local Similarity 54.4%; Pred. No. 1e-22;  
 Matches 267; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

Qy 126 CTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTT 185  
 Db 104 CTCCTGACCAAAACGGGCTTCCAGTTCTACTACCTGCCACCGCTCTACATCCTCGTCT 163  
 Qy 186 TATCCTAGGCGTGCACATAATGGCACTGTCTTGTGGCACTCTCTGGGCGCAACCAAGCG 245  
 Db 164 CATCACGGGTTCTTGGGCAACAGCGTGGCCATCTGGATGTTCTTCCACATGCGGCC 223  
 Qy 246 CTGAGAGTGTGCCACCACTATCTGGTGAACCTGATGTGGCGGCACTTATGTGCT 305  
 Db 224 GTGAGCGGCATCTCGGTGTACATGTTCAACCTGGCTCTGGCCGCACTTCTGTATGCT 283  
 Qy 306 AT---TGCCCTTCTCATCATCATCTACTACTAGTACAGAGTGGCCCTTCGGGAGCT 362  
 Db 284 CACGCTGCGCCGCCCTCATCTTCTACTTCAACAAACCGCACTGGATCTTCGGGAGCT 343  
 Qy 363 GCTCTGAAGCTGTGCACTCTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCT 422  
 Db 344 CATGTGAAGCTGCAGAGGTTTCATTTTCCAGTGAACCTCTACGGCAGCATCTGTTCT 403  
 Qy 423 GACCTGATCTCTGTGACCAAGTCTCTAGTGTGTGGCCACCACTGTGTGCTGCGCCTA 482  
 Db 404 CACGTGATAGCTGTGCACAGTACACGGGGTGTGTGCACCCGCTGAAGTCTGCTGGGAG 463  
 Qy 483 CCGAACCCGCAAGGATGCTGCTGGTGGGCAACCACTGACCACTGGGCGCTGTGCTTCCA 542  
 Db 464 GCTGAAGAAGAAGAACGCGGTGTACGTGCTGTGTGGCGCTGTGTGGTGGCGCT 523  
 Qy 543 GCTGCTGCCGCACTGGGCTTCTCCACACGCACTACATCAATGGCCAGATGATCTGTA 602  
 Db 524 CATCGCGCCCATCTCTTCTACTCAGGACGNNNGTGAGGAGGAACAAACCATCACGTG 583  
 Qy 603 TGACATGACCA 613  
 Db 584 CTACGACACCA 594

RESULT 7  
 BQ33118

LOCUS BI833118 851 bp mRNA linear EST 04-OCT-2001  
 DEFINITION 603090834F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5229633 5',  
 mRNA sequence.

ACCESSION BI833118  
 VERSION BI833118.1 GI:15944668  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 851)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1578 row: c column: 10  
 High quality sequence stop: 847.

## FEATURES

source  
 1. .851  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5229633"  
 /clone\_lib="NIH\_MGC\_120"  
 /lab\_host="DH10B"  
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."  
 BASE COUNT 127 a 311 c 231 g 182 t  
 ORIGIN

Query Match 11.2%; Score 121.6; DB 13; Length 851;  
 Best Local Similarity 53.3%; Pred. No. 3.2e-22;  
 Matches 303; Conservative 0; Mismatches 259; Indels 6; Gaps 2;  
 QY 126 CTGCCAGTTCACAGAGTACAGCAAGTCTACCTCTCCCTGCGCTACAGTATCATCTT 185  
 Db 157 CTGTGTCTACCGGACGAGTCTCAACCACTGTCTGCCACCTGTATTGGCGGTGCT 216  
 QY 186 TATCCTAGGGTGGCCTAAATGCACCTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245  
 Db 217 GCGCGTGGCTGCGCTGAACATCTGTGTATACCACTATGACGATCTGCGAGTCCCGCGGGC 276  
 QY 246 CTGGAGTGTGGCCACCACTCTGTGTGAACCTGATGTGGCGGACCTGCTTTATGTGCT 305  
 Db 277 CCTGCCCGGACGCGGTGTACACCTTAAACCTTGTCTGTGCTGACCTGCTATATGCTG 336  
 QY 306 AT---TGCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCT 362  
 Db 337 CTCCTGCGGCTGTCTCATCTACAACTATGCCCAAGGTGATCACTGGCCCTTTGGCGACT 396  
 QY 363 GCTCTGAAGTGTGGTGCCTCTCTCTATATCAACTTTACGGCAGCATCTGCTGCT 422  
 Db 397 CGCTGCGGCTGTGCGCTTCTCTCTATGCCAACTGTCACGCGCAGCATCTTCTCT 456  
 QY 423 GACCTGATCTGTGTGACCAAGTTCCTAGTGTGTGCCACCACTGTTGCGTGGCCTA 482  
 Db 457 CACCTGATCATAGTTCACGCGCTACCTGGGCATCTGCCACCGCTGGCCCGCTGGCA 516  
 QY 483 CC---GGACCGCAGGATGCTGCTGGGCGGACCAAGCAGCAGTGGCCCTTGGTGGTCT 539  
 Db 517 ACCTGGGGGCGCGGGCTGCTGGCTAGTGTGTAGCGCGTGTGGTGGCGGACAA 576  
 QY 540 CCAGCTGTGCGCCACACTGGCTCTCCACACGAGCATACATCAATGGCCAGATGATCTG 599  
 Db 577 CCATGCTGCGCCACAGCCATCTCTGCTGCGCAGAGGATCCAGGTAACCCCATCTGCTG 636  
 QY 600 GTATGATGATGACCGCAAGAGAATTTGATCGGCTTTTTCCTACGCGATAGTCTGAC 659  
 Db 637 CTATGACCTCAGCGCGCTGCGCTGGCCACCACTATATGCTTATGCTGCTGCTGCTCAC 696  
 QY 660 ATTGCTGGCTTTCTTTCCCTCTCTTGT 687  
 Db 697 TGTCTGCGCTTCTGCTGCGCTTTGCT 724

## RESULT 8

AL561864/c 928 bp mRNA linear EST 16-FEB-2001  
 LOCUS LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB006YB03 3  
 DEFINITION prime, mRNA sequence.  
 AL561864  
 ACCESSION AL561864.1 GI:12909717  
 VERSION  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 928)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

source  
 1. .928  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB006YB03"  
 /clone\_lib="LTI\_NFL004\_NBC2"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"  
 BASE COUNT 188 a 247 c 331 g 159 t 3 others  
 ORIGIN

Query Match 11.1%; Score 120.6; DB 9; Length 928;  
 Best Local Similarity 52.9%; Pred. No. 6.2e-22;  
 Matches 324; Conservative 1; Mismatches 281; Indels 6; Gaps 3;  
 QY 366 CTGCAAGCTGTGCACTTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTGCTGAC 425  
 Db 927 CTGCGGCTGTGCTGCTCTCTTATGCCAACCTGCAAGCAGCAGCATCTCTTCTCTCAC 868  
 QY 426 CTGCATCTCTGTGACCAAGTTCCTAGGTGTGGCACCACCTGTTGCTGCTGCTGCTGCT 484  
 Db 867 CTGCATACAGTTCACGCGCTACCTGGGCATCTGCCACCGCTGGCCCGCTGGCAACAG 808  
 QY 485 --GGACCGCAGGATGCTGCTGGGCACCAAGCAGCAGTGGGCGCTGGTGGTCTCTCA 542  
 Db 807 TGGGGGCGCGGCTGCTGCTGCTAGTGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 748  
 QY 543 GCTGCTGCGCCACACTGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGTA 602  
 Db 747 GTGCTGCGCCACAGCATCTCTGCTGCTGCCACAGGATCCAGTAAACCGCACTGTCTGCTA 688  
 QY 603 TGACATGACCAAGCAAGAAATTTTGATCGGCTTTTTCCTACGCGCATAGTCTGACATT 662  
 Db 687 TGACCTCAGCCCGGCTGCGCTGGCCACCACTATATGCTGCTGCTGCTGCTGCTGCTGCT 628  
 QY 663 GTCTGGCTTCT--TTCCCTCTTGGTCAATTTTGGTGTGCTATTTCACATGATGGTGGT 720  
 Db 627 CATCGGCTTCTGCTGCGCTTTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568  
 QY 721 CCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGC-CGAGCCAGGTCTCA 779









```

VERSION      AL318925.1  GI:9551809
KEYWORDS     GSS: genome survey sequence.
SOURCE       Tetraodon nigroviridis.
ORGANISM     Tetraodon nigroviridis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthopterygii; Percomorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodon.
REFERENCE    1 (bases 1 to 1101)
AUTHORS      Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
              Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
              Saurin, W. and Weissenbach, J.
TITLE        Estimate of human gene number provided by genome-wide analysis
              using Tetraodon nigroviridis DNA sequence
JOURNAL      Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE      20296633
PUBMED      10835645
REFERENCE    2 (bases 1 to 1101)
AUTHORS      Crolius, H.R., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C.,
              Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W.,
              Bernot, A. and Weissenbach, J.
TITLE        Characterization and repeat analysis of the compact genome of the
              freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Genome Res. 10 (7), 939-949 (2000)
PUBMED      20359837
REFERENCE    3 (bases 1 to 1101)
AUTHORS      Direct Submission
TITLE        Submitted (12-APR-2000)
JOURNAL      This sequence is a single read and was generated as part of a large
              scale clone-end sequencing project of the Tetraodon nigroviridis
              genome. For more information, please take a look at
              http://www.genoscope.cns.fr/tetraodon.
FEATURES     Location/Qualifiers
              1. 1101
              /organism="Tetraodon nigroviridis"
              /db_xref="taxon:99883"
              /clone="020M21"
              /clone_lib="A"
              /note="Genoscope sequence ID : COA0020AG11C1-end : T7"
BASE COUNT   305 a 216 c 280 g 284 t 16 others
ORIGIN
Query Match 11.1%; Score 119.6; DB 17; Length 1101;
Best Local Similarity 53.3%; Pred. NO. 1.2e-21;
Matches 288; Conservative 0; Mismatches 247; Indels 5; Gaps 2;
QY 116 ACAAGGTCTCTGCAGTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGGCTACA 175
DB 537 ATCAACTACAGTGCCTTTGACGAGGACTTCAATATACATCTCTCGCCGCTGCTATG 478
176 GTATCATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCC 235
DB 477 GCCTGGTCTTGTCTTCGGCTTCATCTCAAGCCACAGCCCTGTGGCTCTTCCTGA---421
QY 236 AAACAAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGTAGTGGCCGACCTGC 295
DB 420 AGATGGCGCCCTGGAACCCCAACACCGCTCTTCTTGTCCACCTCGCCCTCTCCGACTTC 361
QY 296 TTATGTGCTATTGGCTTCCTCATCATCACTACTACTAGATACAGGTG--GCCCTT 353
DB 360 TCTACATCTGTGCTGGCCACCTCTATCTACTATATAGCCCAACCAACGCGCACTGGCCTT 301
QY 354 CGGGAGCTGCTGCAAGCTGGTGCATCTTCTGTCTATATCAACCTTTACGGCAGCAT 413
DB 300 CGGCGTGGCAGCTGCAAAATTTGGCTTCTCTTCTTCTACGTCACCTCTACTGCAGCAT 241
QY 414 CTTGCTGTGACCTGCATCTCTGTGCACAGTTCCTAGTGTGTGCCACCCACTGTGTC 473
DB 240 CCTCTTCCTACCTGCATCAGCGTGTACCGTGTACGTGGGCATCTGCCACCCGATTAAGGC 181
QY 474 GCTGCCCTACCGGACCCGAGCATGCCTGGCTGGCGCCACCAACCACTGCGGCCCTGTGT 533

```

```

DB 180 GCTGAATCTGTTGAANCCCGGANATTCCTACCTGTTGCGCTCGGTGGGCTGTGCT 121
QY 534 GGTCTCTCCAGCTGCTGCCACACACTGGCTTCTCCACAGGACTACATCAATGGCCAGAT 593
DB 120 CATAGTCTGTCTGGTGGCTTAACCTCATCTTTGTACCACCTCCAGGCGGGAACAATCNC 61
QY 594 GATCTGGTATGATGACACGACCAAGAAATTTTGTATCGCTTTTTCCTAGGCATAGT 653
DB 60 CCTGTGCCATGATACACACNCCGCGAGTTATTCTGCAATATGTAATAATTACTGCNNCT 1

RESULT 13
LOCUS      AL521440
DEFINITION AL521440 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB001YJ16.5
              prime, mRNA sequence.
ACCESSION  AL521440
VERSION     AL521440.1  GI:12784933
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 955)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 Evry cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES    Location/Qualifiers
              1. 955
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="CS0DB001XJ16"
              /clone_lib="LTI_NFL004_NBC2"
              /sex="male"
              /tissue_type="neuroblastoma cells"
              /lab_host="DH10B"
              /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
              was primed with a NotI-oligo(dT) primer. Five prime end
              enriched, double-stranded cDNA was digested with NotI and
              cloned into the NotI and EcoRV sites of the pCMVSPORT 6
              vector. Library was normalized. Library was constructed
              by Life Technologies. Contact : Feng Liang Life
              Technologies, a division of Invitrogen 9800 Medical Center
              Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
              8371 Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com"
BASE COUNT 160 a 400 c 159 g 209 t 27 others
ORIGIN
Query Match 10.5%; Score 113.8; DB 9; Length 955;
Best Local Similarity 49.7%; Pred. NO. 4.5e-20;
Matches 283; Conservative 20; Mismatches 260; Indels 6; Gaps 2;
QY 126 CTGCAGTCTTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCTCAGATCATCTT 185
DB 285 CTGTCTCTACCGGAGAACTTCAAGCAACTGCTGCTGCCACCTGTGTATTCGCGGTGCT 344
QY 186 TATCCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAACCAAGG 245
DB 345 GCGGCTGGCCCTGCCGCTGAACATCTGTCTATTACCCAGATCTGCACGTCGCCGGGC 404
QY 246 CTGGAGCTGTCCACCTATCTGTGAACCTGTATGTGGCCGCTCTTATGTGCT 305
DB 405 CTGACCCGCGAGCGCGGTGACACCTTAACCTTCTGCTGGCTGACCTGCTATATGCTG 464
QY 306 AT---TGGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGSAGCT 362
DB 465 CTCCTGCCCCCTGCTCATCTACAACTATGCCCAAGGTGATCACTGGCCCTTTGGCGANTT 524

```

```

QY 363 GCTCTGCAAGCTGGTGCATCTCTCTCTCTATATCAACCTTTACGGCAGCATCTGCTGCT 422
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 525 CGCTGCMSCCTGGTCCGCTTCTCTCTATGCAACCTGCAAGCGCAGCATCTCTTCTCT 584
QY 423 GACCTGCATCTCTGTCACCAAGTTCCTAGTGTGTGCGCACCACCTGTTGCGTGCCTTA 482
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 585 CACCTGTCATCTCTGACCGCTACCTGGGCATCTCGACCCGCTGGCACCCCTGGCACAA 644
QY 483 CC---GGACCCGAGCATGCTGGCTGGGCACACGACCTGGGCCCTGGTGTCTCT 539
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 645 ACSTGSSSCSCSSCTSSCTSTAGTGTSTAGCCCTSTSCCTCCCTCTACAA 704
QY 540 CCAGCTGTGCCACACTGGCTTCTCCACACGAGTACATCAATGCGCAGATGATCTG 599
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 705 CCACCTCCCTCCACACCATNTTBCCTCCACACATMCACCCVAAACCCACTCTCTC 764
QY 600 GTATGACATGACCAAGCAATTTTGTATGGCTTTTGTGCTAGCGGATAGTCTGAC 659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 765 CTATCAACTACACCCCTCCCTCCACACCACTATATCCCTATCCCATCCCTCTCAC 824
QY 660 ATTGTGCTGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 688
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 825 TCTCATYCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 853

RESULT 14
LOCUS B1819396
DEFINITION 603034571F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:517561 5',
      mRNA sequence.
ACCESSION B1819396
VERSION B1819396.1 GI:15930946
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL 1 (bases 1 to 898)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs@mail.nih.gov
      Tissue Procurement: Life Technologies, Inc.
      cDNA Library Preparation: Life Technologies, Inc.
      cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LLAM1437 row: n column: 18
      High quality sequence stop: 884.
      Location/Qualifiers
        1..898
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="IMAGE:517561"
          /clone_lib="NIH_MGC_115"
          /lab_host="DH10B"
          /note="Organ: pooled brain, lung, testis; Vector:
          pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
          source anonymous pool of 6 male brains, age range 23-27; 1
          male lung, age 27; and 1 male testis, age 69. Library is
          oligo-dT primed and directionally cloned (EcoRV site is
          destroyed upon cloning). Average insert size 1.8 kb,
          insert size range 1-3 kb. Library is normalized and
          enriched for full-length clones and was constructed by C.
          Gruber (Invitrogen). Research Genetics tracking code
          021. Note: this is a NIH_MGC Library."
BASE COUNT 149 a 303 c 255 g 191 t
ORIGIN

```

```

Query Match 10.1%; Score 109.8; DB 13; Length 898;
Best Local Similarity 53.0%; Pred. No. 5.5e-19;
Matches 329; Conservative 0; Mismatches 282; Indels 10; Gaps 4;

QY 126 CTGCCAGTTCTCAGAGAAGTACAAAGCAAGTCTACCTCTCCCTGCGCCTACAGTATCATCTT 185,
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 CTGTGTCTACCGGAGAACTTCAAGCAACTGCTGCTGCCACCTGTGTATTCGGCGGTGCT 301
QY 186 TATCCTTAGGCTGCCACTAAATGGCACTGCTTGTGTGCACTCTCTGGCGCCAAACCAAGCG 245
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 GCGGCTGCGCTGCCCTGAACATCTGTGTCACTTACCAGATCTGCAGCTCCCGCGGGC 361
QY 246 CTGAGCTGTGCGCACCACTATCTGTGAACCTGATGTGTGGCGCACTGCTTTATGTGCT 305
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 CCGTACCGCGCAGCGCGTGTACACCTTAACCTTGTCTGGCTGACCTGATATATGCGCTG 421
QY 306 AT---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAGCT 362
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 CTCCTGCGCCCTGCTCATCTACAACTATGCCAAGGTGATCACTGGCCCTTTGGCGACTT 481
QY 363 GCTCTCAAGCTGCTGCATCTCTCTGTATATCAACCTTTAGCGCAGCATCTGCTGCT 422
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 CGCTGCGCGCTGCTGCTCTCTCTATGCCAACCTGCACGCGCAGCATCTCTTCTCT 541
QY 423 GACCTGCATCTCTGTGCACAGTTCTCTAGTGTGTGCGCACCCACTGTGTGCTGCGCTA 482
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 CACCTGCATCAGCTTCCAGCGCTACCTGGGCATCTGCCACCCGCTGGCCCTGGCACAA 601
QY 483 CC---GGACCCGAGCATGCTGGCTGGGCACACGACCTGGGCCCTGGTGTCTCT 538
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 ACCTGGGGGCGCTTGGCGGCTGCTGCTGTGTGTAGCCGTGTGCTGGCGCTGACAA 661
QY 539 TCCAGCTGTGCGCCACACTGGCTTCTCCACACGCACTACATCAATGGCCAGATGATCT 598
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 CCAGTGCCTGCCACAGCCATCTTCTGCTGCCACAGCATCCAGCGCTAACCGCACTGCT 721
QY 599 GGTATGACATGACCGCAAGAGAAATTTGATCGGCTTTTTC-CCTACGCGATAGTTCTG 657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 GCTATGACCT--CAAGCGCGCTGCTGGGCACCCACTATATGCCCTATGCACTGGCTCTC 779
QY 658 ACATTGCTGCGCTTCTTCCCTCCCTTGGTGTGCTATTCATCATGATGCTGAC 717
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 780 ACTGTCATCGCTTCTGTCGCTTGGCTGCTGCTGTGCTGTGCTGCTGCTGCTGCTGCT 839
QY 718 GAGCCTGATCAAGCCAGAGGA 738
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 840 GTCCCGCAGGATGCCGGAGA 860

RESULT 15
LOCUS AL525099
DEFINITION AL525099 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC0051M02 5
      prime, mRNA sequence.
ACCESSION AL525099
VERSION AL525099.1 GI:12788592
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 884)
JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
      Unpublished (2001)
      Contact: Genoscope
      Genoscope - Centre National de Sequencage
      Bp 191 91006 EVRY cedex - France
      Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
      Location/Qualifiers
        1..884
          /organism="Homo sapiens"
          /db_xref="taxon:9606"

```

